The correct dates will be entered by the editor

Analysis of Fractals, Image Compression, Entropy Encoding, Karhunen-Loève Transforms

Palle E. T. Jorgensen · **Myung-Sin Song**

Abstract The purpose of this paper is to show that algorithms in a diverse set of applications may be cast in the context of relations on a finite set of operators in Hilbert space. The Cuntz relations for a finite set of isometries form a prototype of these relations. Such applications as entropy encoding, analysis of correlation matrices (Karhunen-Loève), fractional Brownian motion, and fractals more generally, admit multi-scales. In signal/image processing, this may be implemented with recursive algorithms using subdivisions of frequency-bands; and in fractals with scale similarity. In Karhunen-Loève analysis, we introduce a diagionalization procedure; and we show how the Hilbert space formulation offers a unifying approach; as well as suggesting new results.

1 Introduction

1.1 General Background and motivation

This paper covers topics from two general areas which we will refer to as: (a) mathematics of signals and image processing; and (b) techniques from theory of operators and Hilbert space. A main novelty is a host of interconnections between the two subjects.

Palle E. T. Jorgensen
Department of Mathematics, The University of Iowa, Iowa City, IA52242, USA
Tel.: +1-319-335-0782
Fax: +1-319-335-0627
E-mail: jorgen@math.uiowa.edu

Myung-Sin Song
Department of Mathematics and Statistics, Southern Illinois University Edwardsville, Edwardsville, IL62026, USA
Tel.: +1-618-650-2580
Fax: +1-618-650-3771
E-mail: msong@siue.edu

Our intention is to turn some of applied problems (from engineering) into rigorous mathematics. Often we encounter various engineering problems initially in a form that has not yet been converted into mathematics, and a good part of our work is then to do that.

1.1.1 (a) Signal and Image Processing

Wavelet theory stands on the crossroads of signal processing, operator theory, and harmonic analysis. It is concerned with the mathematical tools of digitizing continuous data with a view to storage and compression; and with a synthesis process, serving to recreate a desired picture (or time signal) from the stored data. The algorithms involved go under the name of filter banks (see Figure 1), and their remarkable efficiency derives in part from the use of hidden self-similarity in the data or images that are analyzed. Self-similarity is built into wavelets too as they are intrinsically defined using dynamical and iterated function systems./ [JS07b], [JS07a], [DJ06a]

A systematic study of bases in Hilbert spaces built on fractals suggests a common theme: A hierarchical multiscale structure. A well-known instance of the self-similarity is reflected in the scaling rules from wavelet theory [Dau92], [Dau93]. See (1) below. The best known instances of this are the dyadic wavelets in $L^2(\mathbb{R})$. They are built by two functions φ and ψ; subject to the relation

$$\varphi(x) = 2 \sum h_n \varphi(2x - n), \text{ and } \psi(x) = 2 \sum g_n \varphi(2x - n). \tag{1}$$

where (h_n) and (g_n) are fixed and carefully chosen sequences. The function φ is called the scaling function or the father function, and ψ is called the mother function. The best known choice of pairs of filter coefficients (h_n), (g_n) is the following: Pick $(h_n) \subset \mathbb{R}$ subject to the two conditions $\sum_{n\in\mathbb{Z}} h_n = 1$ and $\sum_{n\in\mathbb{Z}} h_n h_{n+2l} = \frac{1}{2}\delta_{0,l}$. Then set $g_n := (-1)^n h_{1-n}$, $n \in \mathbb{Z}$. (The convention is that (h_n) is 0 outside some specified range.) With careful choice of the two sequences, the associated double indexed family $\psi_{jk}(x) = 2^{i/2}\psi(2^j x - k)$, $j, k, \in \mathbb{Z}$ will be a wavelet basis for $L^2(\mathbb{R})$. This is the best known wavelet construction also known by the name multi-resolution analysis (MRA). The reason for this is that the father function φ generates a subspace V_0 of $L^2(\mathbb{R})$ which represents a choice of resolution for wavelet decomposition. In the technical terms, see below, $(\psi_{j,k})$ turn into a Parseval frame in the Hilbert space $L^2(\mathbb{R})$.

One of the advantages of the MRA-method is that it offers a link between the continuous and the discrete wavelet transforms. Further, its use of binary-tree algorithms is practical for computations, and for engineers doing signal and image processing; see [Jor06].

Definition 1 Take V_0 to be the closed span of all the translates $(\varphi(\cdot - k))$, $k \in \mathbb{Z}$ in $L^2(\mathbb{R})$. From (1), it follows that the scaling operator $Uf(x) = 2^{-1/2}f(\frac{x}{2})$ maps the space V_0 into itself; and that $U\psi \in V_0$.

With suitable modification this idea also works for wavelet bases in $L^2(\mathbb{R}^d)$, and in Hilbert spaces built on fractals [DJ06a].

In fact, this very same idea has uses outside the case of affine wavelets. To see this note that the scaling operator U above could be realized by any other transformation possibly unrelated to $x \to \frac{x}{2}$ which was used above.

Selfsimilarity. For Julia sets X in complex analysis for example, U could be implemented by a rational function $z \mapsto r(z)$. Recall that when r is given, X will

be a compact subset of $\mathbb{C}$ which is determined by the dynamics of $r^n = \underbrace{r \circ \cdots \circ r}_{n \text{ times}}$.

Specifically,

$$\mathbb{C} \backslash X = \cup \{\mathcal{O} | \mathcal{O} \text{ open}, (r^{(n)} | \mathcal{O}) \text{ is normal}\}. \tag{2}$$

We want to show that these non-linear fractals are related to more traditional wavelets, i.e., those of $L^2(\mathbb{R}^d)$. We want to extend the $\mathbb{R}^d$-analysis both to fractals and to discrete hierarchical models.

In this paper, two computational features will be addressed: (a) Approximation of the father or mother functions by subdivision schemes, and (b) matrix formulas for the wavelet coefficients. A variety of data will be considered; typically for fractals, L^2-convergence is more restrictive than is the case for $L^2(\mathbb{R}^d)$-wavelets.

This makes wavelets closely related to fractals and fractal processes [Bar02], [Bar96], [BB97], [Bar06], [BEH89], [BHS05], [DL06], [DL92], [BY06], [Bar98], [Bra06], [DRS07]. Investigation of this relation would have theoretical and practical potential and thus is becoming a subject of growing interest both in and outside mathematics. It has been recently shown (by Palle Jorgensen [Jor06], Ola Bratteli [BJ02], David Larson, X. Dai [DL98], [DLS98], [HL00], [LSST06], [LTW06], [Lar07] and others) that a unifying approach to wavelets, dynamical systems, iterated function systems, self-similarity and fractals may be based on the systematic use of operator analysis and representation theory. See Figure 1 below.

1.2 Operator Theoretic Models

Motivated by hierarchical models and multiscaling, we study operators of multiplication, and dilations, and more general weighted composition operators. In these models, scaling is implemented by non-linear and non-invertible transformations. This in turn generalizes affine transformations of variables from wavelet analysis and analysis on affine fractals.

The properties of dynamical and iterated function systems, defined by these transformations, govern the spectral properties and corresponding subspace decompositions. Operator algebras generated by such operators are concrete representations of generalized crossed product type algebras and C^*-algebras defined via various kinds and generalizations of covariance relations. See Figure 1. Invariant sets and measures for the dynamical and for the iterated function systems will be orbits or attractors; and invariant or quasi-invariant measures on them are directly linked to such operator algebras.

The interplay between dynamical and iterated function systems and actions of groups and semigroups on one side, and operator algebras on the other side, yield new results and methods for wavelets and fractal analysis and geometry.

There are many applications of the above mentioned methods and analysis to engineering and physics problems, as well as possibilities to gain insight into numerical analysis of corresponding applications.

Tools from diverse areas of analysis, as well as from dynamical systems and operator theory, merge into the research on wavelet analysis.

Operator methods enter in that wavelets, signals and information may be realized as vectors in a real or complex Hilbert space. For the purpose of transmission, these vectors are encoded in for example a set of linear coefficients. In the case of images, including fractal images, this may be worked out using wavelet and filter functions, for

example corresponding to ordinary Cantor fractal subsets of $\mathbb{R}$, as well as for fractal measure spaces of Sierpinski Gasket fractals.

Fractals, finitely summable infinite trees, and Sierpinski gaskets fit naturally within this framework: Use of spectral triples [Con94] in the computation of geodesic distances and of Minkowski dimension. Furthermore, in the case of the Sierpinski gasket, the associated Dixmier-type trace coincides with the normalized Hausdorff measure.

Other results include applications of a theory of projections, of decompositions of representations, and a theory of operator-valued frames.

1.2.1 (b) Operators and Hilbert space

Operator algebra constructions of covariant representations are used in the analysis of orthogonality in wavelet theory, in the construction of super-wavelets, and orthogonal Fourier bases for affine fractal measures. See e.g. [DJ06b], [DJ06c], [Jor06], [JS07b], [JS07a], [JP98c], [JP98a], [JP98b], [Kig01], [Kig03], [LP06], [OS07], [Str06b], [Str06a].

In motivating our approach, we take advantage of the fact that Hilbert space and operator theory form the common language of both quantum mechanics and of signal/image processing. Recall that in quantum mechanics, (pure) states as mathematical entities "are" one-dimensional subspaces in complex Hilbert space $\mathcal{H}$, so we may represent them by vectors of norm one. Observables "are" selfadjoint operators in $\mathcal{H}$, and the measurement problem entails von Neumann's spectral theorem applied to the operators.

In signal processing, time-series, or matrices of pixel numbers may similarly be realized by vectors in Hilbert space $\mathcal{H}$. The probability distribution of quantum mechanical observables (state space $\mathcal{H}$) may be represented by choices of orthonormal bases (ONBs) in $\mathcal{H}$ in the usual way (see e.g., [Jor06]). In signal/image processing, because of aliasing, it is practical to generalize the notion of ONB, and this takes the form of what is "a system of frame vectors"; see [Chr03] and Definition 4.

But even von Neumann's measurement problem, viewing experimental data as part of a bigger environment (see e.g., [D'A97], [Wan04], [Eld03]) leads to basis notions more general than ONBs. They are commonly known as Positive Operator Valued Measures (POVMs) and frames. We will examine the common ground between the two seemingly different uses of operator theory in the separate applications.

We show that parallel problems in quantum mechanics and in signal processing entail the choice of "good" orthonormal bases (ONBs). One particular such ONB goes under the name "the Karhunen-Loève basis." It is optimal in three ways, and we will outline a number of applications.

Our problems are motivated by consideration of the optimal choices of bases for certain analogue-to-digital (A-to-D) problems we encountered in the use of wavelet bases in image-processing (see [GWE04], [SCE01], [Use01], [Wal99]); but certain of our considerations have an operator theoretic flavor which we wish to isolate, as it seems to be of independent interest.

Several reasons: Firstly our Hilbert space results seem to be of general interest outside the particular applied context where we encountered them. And secondly, we feel that our more abstract results might inspire workers in operator theory and approximation theory.

2 Entropy Encoding and Applications

In the following discussion, we focus on the following three topics: Entropy Encoding, Karhunen-Loève Transforms and Digital Image Compression.

2.1 Entropy Encoding

The following is an elementary example of Shannon-Fano entropy encoding on a text: letters-to-codes:

Letter	Probability	code
a	0.3	00
e	0.2	01
f	0.2	100
q	0.2	101
r	0.1	110

While this is an oversimplification, it is nonetheless a key idea used in more realistic algorithms:

- In a given text, list all letters in decreasing order of their probabilities.
- Divide the list into two parts with approximately equal probability (i.e., by the median, the total probability of each part is approximately 0.5).
- For the letters in the first part, start the code with a 0 bit, and for those in the second part with a 1.
- Recursively continue until each subdivision is left with just one letter [Bel90].

Note that the divisions of the list are following a binary tree-rule. The initial important uses of encoding were to texts and to signals. Much more recent uses are to a host of big data sets such as data on the color images. These uses are: quantization, entropy encoding. As a result, the mathematics of encoding has seen a recent revival.

While entropy encoding is popular in engineering, [SCE01], [Use01], [DVDD98], the choices made in signal processing are often more by trial and error than by theory. Reviewing the literature, we found that the mathematical foundation of the current use of entropy in encoding deserves closer attention.

2.1.1 Benefits of Entropy Encoding

An important part of image processing is digital compression, see Figure 2. This process begins with quantization.

One might think that the quantization step would suffice for compression, this is not so. It is true that quantization does compress data tremendously. After the quantization step, many of the pixel values are either eliminated or replaced with other suitable values. However, those pixel values are still represented with either 8 or 16 bits. So we aim to minimize the number of bits used by means of entropy encoding. Karhunen-Loève transforms or PCAs make it possible to represent each pixel on the digital image with the least-bit representation according to their probability thus yielding a lossless optimized representation, using least amount of memory.

Probability. Image pixels are realizations of independent, identically distributed (i.i.d.) random varibles. The probability density function for these random variables is

determined empirically by counting frequencies of occurrence of various pixel values. Entropy coders provide efficient bit representations of pixel values by assigning short bit strings to frequently occurring values and longer strings to less frequently occurring values. Treating pixel values as i.i.d. random variables, and entropy coding them, in practice yields a modest decrease in storage requirements; and this process is lossless.

In digital image compression, after the quantization step (see Figure 2), entropy encoding is performed on a particular image for more efficient storage (i.e., less storage memory). When an image is to be stored, we need either 8 bits or 16 bits to store a pixel. With efficient entropy encoding, we can instead use a smaller number of bits to represent a pixel in an image, resulting in less memory used to store or even transmit an image.

This paper extends the work we started in [JS07b]. The basic principle in signals and images is the breaking up of input into non-overlapping parts. By this we mean, for example, into uncorrelated bands or into orthogonal components. In our work, we use Hilbert space in describing orthogonality relations. Since signals are quantized by size of energy, it is useful to encode data in Hilbert space in such a way that the processing is done by isometries. The requirement that the bands are orthogonal amounts to asking the range of the isometries to be orthogonal spaces. Asking perfect reconstruction then amounts to a decomposition of the identity operator. These exact requirements are met by representations of relations known in Operator Algebra theory as the Cuntz relations; see [Cun77, ?] and Figure 1.

Definition 2 For every finite n, a representation of the Cuntz algebra $\mathcal{O}_n$ is a system of isometries $S_i : \mathcal{H} \to \mathcal{H}$ such that

(a) $S_i^* S_j = \delta_{ij} I$; orthogonality, and
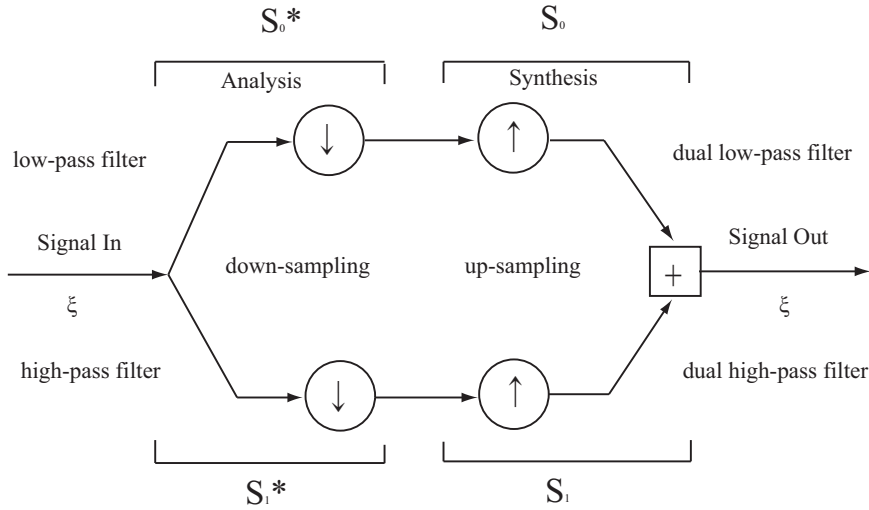(b) $\sum_i S_i S_i^* = I$ (perfect reconstruction).

See Figure 1 below.

Fig. 1 Examples of operators in signal image processing.

3 Slanted Matrix Representations, Frame Estimates, and Computations

In this section we prove mathematical theorems supporting the applications outlined above: The operators in Fig 1 have slanted matrix representations determined by the masking sequences (h_n) and (g_n), and with the slanting changing from one operator S to the corresponding adjoint operator S^*. We then show how frame estimates from Definition 4 are preserved under filtering with our S-systems, i.e., with the slanted matrices that realize the Cuntz relations in (a) and (b) above, The slanted matrix representation is what make computations fast. The slanting is such that repeated matrix operations in the processing make for more sparse matrices, and hence for a smaller number of computational steps in digital program operations for image processing.

We begin by introducing the Cuntz operators S . The two operators come from the two masking sequences (h_n) and (g_n) , also called filter-coefficients, also called low-pass and high-pass filters.

Definition 3 If $(h_n)_{n\in\mathbb{Z}}$ is a double infinite sequence of complex numbers, i.e., $h_n \in \mathbb{C}$, for all $n \in \mathbb{Z}$; set

$$(S_0 x)(m) = \sqrt{2} \sum_{n\in\mathbb{Z}} h_{m-2n} x(n) \tag{3}$$

and adjoint

$$(S_0^* x)(m) = \sqrt{2} \sum_{n\in\mathbb{Z}} \bar{h}_{n-2m} x(n); \text{ for all } m \in \mathbb{Z}. \tag{4}$$

Then

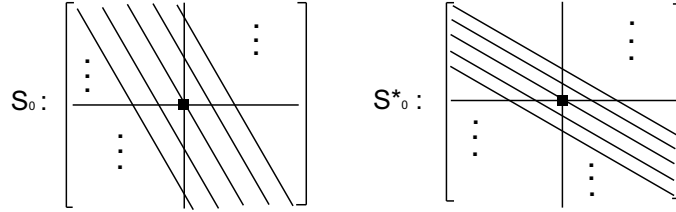(a) The $\infty \times \infty$ matrix representations (3) and (4) have the following slanted forms

Fig. 2 S_0 and S_0^*.

(b) The set of non-zero numbers in $(h_n)_{n\in\mathbb{Z}}$ is finite if and only if the two matrices in Fig are *banded*.
(c) Relative to the inner product

$$\langle x|y\rangle_{l^2} := \sum_{n\in F} \bar{x}_n y_n \text{ in } l^2$$

(i.e., conjugate-linear in the first variable), the operator S_0 is *isometric* if and only if

$$\sum_{n\in F} \bar{h}_n h_{n+2p} = \frac{1}{2}\delta_{0,p}, \text{ for all } p \in \mathbb{Z}. \tag{5}$$

(d) If (5) holds, and

$$(S_1 x)(m) = \sqrt{2} \sum_{n \in \mathbb{Z}} g_{m-2n} x(n), \tag{6}$$

then

$$S_0 S_0^* + S_1 S_1^* = I_{l^2} \tag{7}$$

$$S_k^* S_l = \delta_{k,l} I_{l^2} \text{ for all } k, l \in \{0, 1\} \tag{8}$$

(the Cuntz relations) holds for

$$g_n := (-1)^n \bar{g}_{1-n}, \ n \in \mathbb{Z}.$$

Proof By Parseval's identity and Fourier transforms, we have the *isometric* isomorphism $l^2(\mathbb{Z}) \simeq L^2(\mathbb{T})$ where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ is equipped with Haar measure.

Hence the assertions (a)-(d) may be checked instead in the following function representation:

$$f(z) = \sum_{n \in F} x(n) z^n, \tag{9}$$

$$m_0(z) = \sum_{n \in F} h_n z^n, \tag{10}$$

$$m_1(z) = \sum_{n \in F} g_n z^n; \tag{11}$$

setting

$$(S_j f)(z) = m_j(z) f(z^2), \text{ for all } z \in \mathbb{T}, \text{ for all } f \in L^2(\mathbb{T}), \ j = 0, 1. \tag{12}$$

In this form, the reader may check that conditions (a)-(d) are equivalent to the following unitary principle: For almost every $z \in \mathbb{T}$ (relative to Haar measure), we have that the 2×2 matrix

$$U(z) = \begin{pmatrix} m_0(z) & m_0(-z) \\ m_1(z) & m_1(-z) \end{pmatrix} \tag{13}$$

is unitary; i.e., that $U(z)^* = U(z)^{-1}$, almost every $z \in \mathbb{T}$, where $(U^*)_{k,l} := \bar{U}_{l,k}$ denotes the adjoint matrix.

Suppose A is an index set and $(v_\alpha)_{\alpha \in A} \subset l^2(\mathbb{Z})$ is a system of vectors subject to the frame bound $(B < \infty)$

$$\sum_{\alpha \in A} |\langle v_\alpha | x \rangle_{l^2}|^2 \leq B \|x\|_{l^2}^2, \text{ for all } x \in l^2; \tag{14}$$

see Definition 4.

Set

$$w_{j,k} := S_0^j S_1 v_\alpha, \ j \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}, \alpha \in A. \tag{15}$$

If the unitarity condition (13) in the lemma is satisfied, then the induced system (15) satisfies the same frame bound (14).

Proof Introducing Dirac's notation for rank-one operators:

$$|u\rangle\langle v||x\rangle = \langle v|x\rangle|u\rangle, \tag{16}$$

we see that (14) is equivalent to the following operator estimate

$$\sum_{\alpha\in A} |v_\alpha\rangle\langle v_\alpha| \le BI_{l^2} \tag{17}$$

where we use the standard ordering of the Hermitian operators, alias matrices.

For the system $(w_{j,\alpha})_{(j,\alpha)\in\mathbb{N}_0\times A}$ in (15), we therefore get

$$\sum_{(j,\alpha)\in\mathbb{N}_0\times A} |w_{j,\alpha}\rangle\langle w_{j,\alpha}| = \sum_j\sum_\alpha S_0^j S_1|v_\alpha\rangle\langle v_\alpha|S_1^* S_0^{*^j}$$

$$\underset{by(17)}{\le} BS_0^j S_1 S_1^* S_0^{*^j}.$$

But the RHS in the last expression is the limit of the finite partial sums

$$\sum_{j=0}^N S_0^j S_1 S_1^* S_0^{*j} = \sum_{j=0}^N S_0^j(I - S_0 S_0^*)S_0^{*j} \tag{18}$$

$$= I_{l^2} - S_0^{N+1} S_0^{*^{N+1}} \tag{19}$$

$$\le I_{l^2} \text{ since} \tag{20}$$

$P_{N+1} := S_0^{N+1} S_0^{*^{N+1}}$ is a projection for all $N\in\mathbb{N}_0$. In fact

$$\cdots \le P_{N+1} \le P_N \le \cdots \le P_1 = S_0 S_0^*;$$

and P_1 denotes the projection onto $S_1 l^2$.

Corollary 1 *The induction* $(v_\alpha) \longrightarrow (w_{j,\alpha})$ *in* (15) *preserves both frame bounds in Definition 4(a) if and only if* $\lim_{N\to\infty} P_N = 0$; *i.e., if and only if* $S_0 : l^2 \to l^2$ *is a shift.*

Proof To verify the lower frame bound for an induced system $(w_{j,\alpha})$ as in (3.16), i.e.,

$$\sum_{j=0}^N S_0^j S_1(S_0^j S_1)^* = I_{l^2} - P_{N+1}.$$

But the proof alone shows that if c_1 is a lower frame-bound for a given system $(v_\alpha)_{\alpha\in A}$, then

$$\sum_{j=0}^N\sum_{\alpha\in A} |\langle w_{j,\alpha}|f\rangle|^2 \ge c_1(\|f\|^2 - \|P_{N+1}f\|^2).$$

Hence c_1 is also a lower frame-bound for $(w_{j,\alpha})$ if and only if $\lim_{N\to\infty} P_N = 0$.

Our previous results show how a finite operator system induces matrix algorithms serving as generalizations of discrete wavelet algorithms.

In the proofs we rely on L^2 Fourier duality: The Hilbert space of L^2-functions on the one-torus $\mathbb{T}$ is isometrically isomorphic with the l^2-sequence space. The isometry being given by Parseval's formula.

In the previous results we restricted to dyadic subdivision, but this restriction is easy to remove. The results carry over to the most general subdivision in any dimension, i.e., for data indexed by a rank d lattice for all values of d.

But there is a more important extension. Stay with the dyadic subdivision, but allow more than two sub-band filters: In our next result we show that the matrix algorithms and the frame extensions carry over mutatis mutandis if the number of filters is arbitrary, and we specify the compatibility conditions.

Theorem 1 *If instead of two operators*

$$S_i : l^2(\mathbb{Z}) \to l^2(\mathbb{Z}), \ i = 0, 1$$

we consider a finite system

$$(S_i x)(m) := \sqrt{2} \sum_{n \in \mathbb{Z}} h^i_{m-2n} x(n), \ i = 0, 1, \cdots, N;$$

or equivalently

$$(S_i f)(z) = \sqrt{2} m_i(z) f(z^2), \ f \in L^2(\mathbb{T})$$

with

$$m_i(z) = \sum_{n \in \mathbb{Z}} h_n^{(i)} z^n, \ z \in \mathbb{T}$$

and

$$f(z) = \sum_{n \in \mathbb{Z}} x(n) z^n, \ z \in \mathbb{T}.$$

then the results alone carry over if and only if one of the following equivalence conditions hold:

$$\left\| \sum_{i=0}^{N} S_i^* f \right\|^2 = \|f\|_{L^2(\mathbb{T})}^2, \ \text{for all } f \in L^2(\mathbb{T}); \tag{21}$$

$$\sum_{i=0}^{N} |m_i(z)|^2 = 1, \ \text{almost all } z \in \mathbb{T} \tag{22}$$

and

$$\sum_{i=0}^{N} \overline{m_i(z)} m_i(-z) = 0, \ \text{almost all } z \in \mathbb{T}. \tag{23}$$

Proof Note that condition (21) states that the vector operator $(S_0^*, S_1^*, \cdots, S_N^*)$ is isometric; i.e., that

$$L^2(\mathbb{T}) \ni f \mapsto \sum_{i=0}^{N} S_i^* f \in L^2(\mathbb{T})$$

is isometric.

The inner product in $L^2(\mathbb{T})$ is defined from the normalized Haar measure μ on $\mathbb{T}$, and we simply compute the LHS in (21):

$$\sum_{i=0}^{N} \left| \frac{1}{\sqrt{2}} \sum_{w^2=z} \overline{m_i(w)} f(w) \right|^2 d\mu(z) = \frac{1}{2} \sum_{i=0}^{N} \int_{\mathbb{T}} \sum_{w} \sum_{w'} m_i(w) \overline{m_i(w')} f(w) \overline{f(w')} d\mu(z),$$
$$w^2 = w'^2 = z$$
$$\underset{by(23)}{=} \frac{1}{2} \int_{\mathbb{T}} \sum_{w^2=z} |m_i(w)|^2 |f(w)|^2 d\mu(z)$$
$$= \int_{\mathbb{T}} \sum_{i=0}^{N} |m_i(z)|^2 |f(z)|^2 d\mu(z)$$
$$\underset{by(22)}{=} |f(z)|^2 d\mu(z) = \|f\|_{L^2(\mathbb{T})}^2.$$

This proves the equivalence of (22) and (23) with (21).

3.0.2 The Karhunen-Loève Transform

The Karhunen-Loève transform is an operator theoretic tool which has proved effective and versatile in the analysis of stochastic processes (X_t). The starting point in this is a spectral analysis of the correlations $E(X_t X_s)$; see 3.2 below. In models, this may represent, for example, correlations of pixel numbers. The K-L analysis involves a variety of choices of bases $(\psi_I(t))$, including wavelet bases, and it leads to a sequence (Z_n) of independent random variables, and an associated K-L expansion of the initial process (X_t).

What motivates transform coders is that our visual surroundings can be closely approximated by smooth surfaces separated by discontinuities that lie along continuous curves. Pixel values are far from independent, particularly pixels which are close together, due to the physical structure beneath real-world images. Given the values of their neighbors, pixels in smooth regions can be predicted with substantial accuracy, so the independent storage of pixels is unnecessary. Exploiting this spatial redundancy (correlation between neighboring pixel values) enables us to acquire a considerable improvement in performance over entropy coding alone. What further motivates this approach is the evidence that the human visual system (HVS) makes use of these types of local smoothness assumptions in low-level visual processing [Dav].

Applying K-L transform to an image yields a set of transform coefficients which are de-correlated, i.e., the first-order linear redundancy in the pixels are eliminated. The quantized linear transform coefficients have the smallest expected mean squared distortion in the K-L transform when the image is recognized as a mean-zero Gaussian process and high resolution quantization is used [Dav].

A number of problems in time series and in stochastic integration suggest adjusting the use standard Brownian motion by its fractional variant (fBM), $X_t^{(H)}$ (t for time); see e.g., [JMR01], [MST99] for details, and [Jor06] for definitions. Here the fraction H is fixed and refers to how the time-variable scales [Jor06] under the substitution $t \mapsto ct$, when c is positive. (The case $H = 1/2$ is standard Brownian motion.) As is known [JMR01], the tool for separating variables in $X_t^{(H)}$ is operator theory (Karhunen-Loève.) Previously, standard Daubechies-Meyer wavelets and fractional integration have been used [MST99]; but to adapt to the H-fractional scaling law, our

fractional wavelet basis [DJ06a] is especially well suited as it has the H-scaling directly built into it. (In some fractal processes, the fraction H is called the Hurst parameter.)

3.1 The Karhunen-Loève Transform Applied to Frames

Multiscale problems such as arise in wavelet theory can often be expressed in terms of unitary operators in Hilbert space. As an example, take dyadic wavelets in the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$: they are the set of unitary operators is $\mathcal{U} := \{U^j T^k | j, k, \in \mathbb{Z}\}$ where U is the scaling operator in Definition 1, and $Tf(x) = f(x-1)$. In this case we look for functions $\psi \in \mathcal{H}$ such that $\{U^j T^k \psi | j, k, \in \mathbb{Z}\}$ is an ONB of $\mathcal{H}$. In many problems involving wavelets and fractals, it is not feasible to look for ONBs, but rather construct frame bases instead. In variety of non-linear applications, one is faced with a set $\mathcal{U} := \{U_a | a \in A\}$ where A is some index set. Then set $C(a,b) := \langle U_a \psi | U_b \psi \rangle$, $a, b, \in A$. In the Hilbert space $l^2(A)$, introduce the following operator $C : (\xi_a) \mapsto (\sum_b C(a,b)\xi_b)$. Note then, that in the case of ONBs the operator C is the identity operator I. In computations, it is often the case that the difference $I - C$ is a compact operator. An application of the spectral theorem will then yield an ONB of eigenvectors $(\mathbf{v}_k)$, such that $(I - C)\mathbf{v}_k = \lambda_k \mathbf{v}_k$, where the eigenvalues λ_k maybe ordered as follows:

$$0 \leq \lambda_1 \leq \lambda_2 \cdots \leq \lambda_k \longrightarrow 1.$$

We want to study spectral theory of fractals with the use of this construction. It is different from but closely related to the Karhunen-Loève technique from information theory. See [Ash90].

3.2 Fractional Brownian Motion (fixed fraction)

The details of this theory has been broken down into fourteen steps which we describe below. While the applications and the framework refers to probability theory and signals, the underlying tools are operator theory.

Rather than diagonalizing using a standard wavelet basis, apply instead our basis functions from [MST99] to the problem of diagonalizing H-fBM, i.e., in the Karhunen-Loève expansions for H-fBM.

The purpose of the following is to outline new *continuous* Karhunen-Loève decompositions (K-L decompositions) and their applications. As it turns out, in the literature to date, rigorous results for K-L decompositions have been restricted to the discrete case. Case in point, for fractional Brownian motion processes (X_t) with t for time, one looks for wavelet bases in the t-variable, and a countably indexed family of Gaussian $N(0,1)$ independent random variables. Then as illustrated in (13) below, the objective is an expansion for processes (X_t) at hand which separates the two variables; t-the time variable, and the other variable in a suitable probability space.

The fraction in fractional Brownian motion fBM is called H for Hurst parameter. The **covariance** function is

(1) $C_H(s,t) := \frac{1}{2}(s^{2H} + t^{2H} - |s-t|^{2H})$, $s, t \in \mathbb{R}$ (s, t for time).
   Notation: (Ω, P) probability space.
(2) $X : \mathbb{R} \times \Omega \to \mathbb{R}$ (or $\mathbb{C}$). The fractional rule for fBM is ($c \in \mathbb{R}_+$ fixed).

(3) $X_{ct} \simeq c^H X_t$ (t: time) where "$\simeq$" means that the two sides have the same distribution: All the joint distributions on both sides in (3) are the same.

If (t_i), $t_i < t_{i+1}$, is a partition, the condition is that $X_{t_{i+1}} - X_{t_i}$, $i = 0, 1, 2, ...$ are correlated according to rule (1) for fixed H. In the special case $H = 1/2$ (Brownian motion), the increments are in fact independent; but not so when $H \neq 1/2$.

The Mean and Covariance conditions are:

(4) $E(X_t) = \int_{\Omega} X_t(\omega) dP(\omega) = 0$,

(5) $E(X_t X_s) = \int X_t(\omega) X_s(\omega) dP(\omega) = C_H(t, s) = \frac{1}{2}(s^{2H} + t^{2H} - |s - t|^{2H})$.

(6) Use operator theory to establish separation of variables formulas for stochastics processes: $X(t, \omega) = \sum_I \psi_I(t) Z_I(\omega)$, $I = (i, j, k)$. We use the triple index notation $I = (i, j, k)$ where i indexes a finite family of mother wavelet functions, and j, k range over $\mathbb{Z}$. The functions under the summation are $\psi_I(t) = \psi_{ijk}(t) = B^j \psi^i(A^j t - k)$, where A and B are numbers depending on the Hurst parameter H. This setup adapts well also to the multivariable case, $t = (t_1, \ldots, t_d)$ where A is then a matrix action on $\mathbb{R}^d$. Further, we demand that $Z_I(\cdot)$ is an independent family of Gaussian random variables with the following properties:

(7) The constants A and B will be chosen such that $\psi_I(\cdot) \in$ Dutkay-Jorgensen wavelet basis, the Dutkay-Jorgensen space-filling wavelets [DJ06a].

The Karhunen-Loève theorem: There then exists Gaussian random variables $Z_I(\cdot)$, and there exists a probability space (Ω, P) such that

(8) $E(Z_I) = 0$ and

(9) $E(Z_I Z_{I'}) = \lambda_I \delta_{I, I'}$, for all I, I'. (Notation: $E(f) := \int f dP$).

So the problem is to find the sequence (λ_I) of eigenvalues. Moreover, there is a measure $\nu = \nu_I$ on $\mathbb{R}$ such that

(10) $\int C_H(s, t) \psi_I(t) d\nu_H(t) = \lambda_I \psi_I(s)$. One then recovers Z_I as follows:

(11) $Z_I(\cdot) = \int_{\mathbb{R}} \psi_I(t) X(t, \cdot) d\nu_H(t)$.

Moreover, the Karhunen-Loève theorem states that then Z_I will be Gaussian, etc. We need to look into the following questions:

1. What can the possibilities be for the measures ν_H?
2. For an admissible ν_H, what is the sequence (λ_I) of eigenvalues?

The Karhunen-Loève theorem is the assertion that the sequence random variables (Z_I), $I \in$ the index set, are Gaussian and independent.

(12) If $E(Z_I^2) = \lambda_I$, we then get (9).

Note that (10) and the Spectral Theorem yield

(13) $C_H(s, t) = \sum_I \lambda_I \psi_I(s) \psi_I(t)$, and

(14) $X(t, \omega) = \sum_I \psi_I(t) Z_I(\omega)$. (See [Jor06] for convergence properties.)

Note that this is consistent with the ansatz (5).

Proof:

$E(X_t X_s) = E(\sum \psi_I(s) Z_I \sum \psi_J(t) Z_J) = \sum_I \sum_J \psi_I(s) \psi_J(t) E(Z_I Z_J)$

$=_{by(9)} \sum_I \psi_I(s) \psi_I(t) \lambda_I =_{by(13)} C_H(s, t)$.

Examples include applications to measurements in Physics of cosmic microwave background radiation (CMB) [APR06], [BKG$^+$], [KP04].

In these applications to CMB, the stochastic process is indexed by points on the 2-sphere (S^2). Some authors have developed a decomposition theory where the functions in the t-variable are generalized wavelet functions called needlets [BKG$^+$].

The basic assumption on these processes are:

(1) rotational invariance, i.e., $E(X_t X_s) = E(X_{R(t)} X_{R(s)})$ for all $s, t \in S^2$, and for all $R \in SO_2$.

(2) X_t is assumed Gaussian.

These two conditions invite use the kind of of representation theory that we have studied [JS07b], [JS07a], [Son07].

Encoding of stochastic processes was studied in [CW84], [Wit04], [Pet86]; and more recently [DS06] it was refined to fractional Brownian motion. It was proved in [DS06] that entropy coding and quantization coincide asymptotically for fractional Brownian motion. This fact will be relevant for our work both on 3.2 above and 3.3 below.

3.3 Applications of Fractional Brownian Motion to Option Pricing

In financial mathematics, certain models use a stochastic differential and integral equation named after Black and Scholes, called the Black-Scholes equation [APS06], [HK06], [JT06], [Jum05], [Mey06], [RM06]. In its simplest form it is

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t).$$

where $S(t)$ denotes the price of a portfolio, μ is a rate and σ is volatility. Its stochastic term in the equation is dW, and W is the standard Brownian motion. The derivation of the Black-Scholes equation depends on two things: first, Ito's lemma and secondly the choice of dW. Since some stock models appear to follow a fractional model, it would be extremely interesting to develop an analogue of the Black-Scholes equation, where the stochastic term in such an equation would be fractional Brownian motion $H \neq 1/2$ as oppose to the standard choice $H = 1/2$ which has been used up to now. By this we mean generalizing the value of the Hurst parameter to $H \neq 1/2$. However, as it stands, the Black-Scholes equation only works if $H = 1/2$. So our proposed research would entail finding a family of equations, one for each value of the Hurst parameter H, such that Black-Scholes comes out as a special case when $H = 1/2$. This generalization is relevant because in realistic models of price fluctuations with the use of stochastic processes, one expects a correlation between time increments for disjoint intervals.

In attacking this problem, we will be using the techniques described above for the study of fractional Brownian motion and of multiscale tools for non-linear dynamics.

3.3.1 Digital Image Compression

In [JS07b], we showed that use of Karhunen-Loève's theorem enables us to pick the best basis for encoding, thus to minimize the entropy and error, to better represent an image for optimal storage or transmission. Here, optimal means it uses least memory space to represent the data; i.e., instead of using 16 bits, use 11 bits. So the best basis found would allow us to better represent the digital image with less storage memory.

The particular orthonomal bases (ONBs) and frames which we use come from the operator theoretic context of the Karhunen-Loève theorem [Ash90]. In approximation problems involving a stochastic component (for example noise removal in time-series or data resulting from image processing) one typically ends up with correlation kernels; in some cases as frame kernels.

Summary of the mathematics used in the various steps of the image compression flow chart in Figure 3 (on next page):
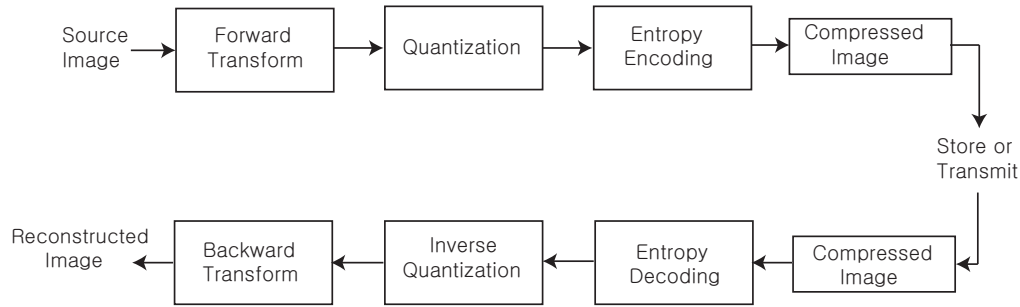
Fig. 3 Outline of the wavelet image compression process. [SCE01]

– At the beginning of the diagram (source image) we will typically input a digital image. Of the possible forward transforms that apply to images, this proposal uses the discrete wavelet transforms.
– The quantization step refers to the following: the output from the transform will be quantized in the sense of economy of processing; for instance, with the thresholding method. After the wavelet forward transform, the image is decomposed into different details at different levels; the thresholding method will eliminate some of these details selectively resulting in lossy compression. In our recent paper [JS07b], we initiated the use of thresholding for exactly this purpose.
– In our approach to image compression the encoding step does the following: with the quantization, the process was lossy where the step is irreversible. With the entropy encoding, if we started off with an 16 bit image we find a better representation meaning we use fewer bits to present the pixel values. This step is lossless. Entropy encoding has been used for long time in information theory, and it has found a recent revival in data compression schemes. The idea is to assign codes to symbols in order to match code lengths with the frequencies of occurrence of the symbols. By entropy we refer to a priori probability distributions of symbols. Entropy encoders compress data by replacing equal length codes with codes where the length of each code would be determined by quantitative measures of entropy. Therefore, the most frequent symbols will be assigned the shortest code. Hence the economy.
There are number of other entropy encoding schemes, Shannon-Fano encoding, Huffman coding, arithmetic coding, etc. [Bel90], [WNC87], [Wit04], [CW84]. These alternative encoding schemes have been successful in signal compression and in simplest static codes, but in the most recent uses of discrete algorithm on images the entropy encoding proved to have practical advantages. The lossless data compression has in common with the discrete wavelet synthesis that exact data can be reconstructed from the compressed versions. Hence, the processing is reversible. Lossless compression is used for example in zip file forward and gzip in Unix/Linux. For images, the formats png and gif also use lossless compression. Examples are executable programs and source codes.
In carrying out compression, one generates the statistical output from the input data, and the next step maps the input data into bit-streams, achieving economy by mapping data into shorter output codes.
– The result of the three steps is the compressed information which could be stored in the memory or transmitted through internet. The technology in both ECW

(Enhanced Compressed Wavelet technology) and JPEG2000 allow for fast loading and display of large image files. JPEG2000 is wavelet-based image compression [Jor06], [JMR01]. There is a different brand new compressing scheme called Enhanced Compressed Wavelet technology (ECW). Both JPEG2000 and ECW are used in reconstruction as well.

3.4 Reconstruction

The methods described above apply to reconstruction questions from signal processing. Below we describe three important instances of this:

(i) In entropy decoding, the fewer bits assigned to represent the digital image are mapped back to the original number of bits without losing any information.
(ii) In inverse quantization, there is not much of recovery to be obtained if thresholding was used for that was a lossy compression. For other quantization methods, this may yield some recovery.
(iii) The backward transform is the wavelet inverse transform where the decomposed image is reconstructed back into an image that is of the same size as the original image in dimension but maybe of smaller size in memory.

Many applications to inverse transforms are done with the use of systems of vectors in Hilbert space which form frames:

Definition 4 (**Frames** [CK04], [Cas06], [CFTW06])

(a) A system $(h_i)_{i \in A} \subset \mathcal{H}$, $\mathcal{H}$ is a Hilbert space, A an index set, is said to be a <u>frame</u> iff there are constants

$$0 < c_1 \leq c_2 < \infty$$

such that

$$c_1 \|f\|^2 \leq \sum_{i \in A} |\langle h_i | f \rangle|^2 \leq c_2 \|f\|^2 \text{ for all } f \in \mathcal{H}.$$

We say that the frame is <u>bounded</u> if there is a constant $c_3 < \infty$ such that $\|h_i\| \leq c_3$ for all $i \in A$.
(b) It is said to be a Riesz basic sequence if

$$c_1 \sum_{i \in A} |\alpha_i|^2 \leq \| \sum_{i \in A} \alpha_i h_i \|^2 \leq c_2 \sum_{i \in A} |\alpha_i|^2 \text{ for all } (\alpha_i) \in l^2.$$

If it is also total in $\mathcal{H}$, then we say that it is a Riesz basis.

3.4.1 The Feichtinger Conjecture [CFTW06]

Every bounded frame is a finite union of Riesz basic sequences.

3.4.2 The Kadison-Singer Conjecture [KS59]

Let $B(\mathcal{H})$ be the algebra of all bounded operators on a separable Hilbert space; and let $\mathcal{D}$ be the subalgebra of operators that are diagonal with respect to an ONB in $\mathcal{H}$. The K-S conjecture states that every pure state on $\mathcal{D}$ has a unique extension to a pure state in $B(\mathcal{H})$.

An exciting new result of Casazza et all [CFTW06] asserts that the Kadison-Singer conjecture [KS59] is equivalent to the Feichtinger conjecture. This is surprising because the K-S conjecture lies at the foundation of quantum physics while the Feichtinger conjecture came from mathematics of signal and image processing. Two new papers by Arveson and Kadison [AK06], [Arv07] add a new dimension to this.

3.5 Connections between the Kadison-Singer Conjecture and Questions in Signal Processing

Our paper is concerned with interplay between the mathematics of these two conjectures, the Kadison-Singer Conjecture and Feichtinger Conjecture. We have worked on a number of specific models in both areas: mathematical physics and rigorous signal processing. The richness of the particular models we have studied seem to suggest some possibilities consistent with negative answers to the two conjectures. (Kadison is inclined to expect a negative answer to the K-S conjecture).

We aim to draw on examples from engineering to explore avenues to a better understanding of two conjectures.

In some cases, parts of the frame vectors fuse (fusion-frames) onto closed subspaces, and we are working with the corresponding family of (orthogonal) projections. Either way, we arrive at a family of selfadjoint positive semidefinite operators in Hilbert space. The particular Hilbert space depends on the application at hand. While the Spectral Theorem does allow us to diagonalize these operators, the direct application of the Spectral Theorem may lead to continuous spectrum which is not directly useful in computations, or it may not be computable by recursive algorithms. As a result we introduce in [JS07b] a weighting of the operator to be analyzed.

The questions we address are optimality of approximation in a variety of ONBs, and the choice of the "best" ONB. Here "best" is given two precise meanings: (1) In the computation of a sequence of approximations to the frame vectors, the error terms must be smallest possible; and similarly (2) we wish to minimize the corresponding sequence of entropy numbers (referring to von Neumann's entropy). In two theorems we make precise an operator theoretic Karhunen-Loève basis, which we show is optimal both in regards to criteria (1) and (2).

We give the problem an operator theoretic formulation. In [JS07b], we introduce the weighting, and we address a third optimality criteria; that of optimal weights: Among all the choices of weights (taking the form of certain discrete probability distributions) turning the initially given operator into trace-class.

3.6 Fractals

Fractals are everywhere in nature and in technology: [Jor06], [Bar96], [BB97], [BEH89], [BHS05], [Dav]. When you look at them in a microscope or in a telescope, you see hidden

patterns as similar repeated structures and features, repeated at different scales. On occasion they are well hidden, for example in huge data sets from the internet. Fractal analysis and data mining are the tools that reveal these features, repeated at varying scales of resolution; and making up fundamental constituents in a yet new and relatively uncharted domain of science.

Intuitively, think of a fractal as reflecting similarity of scales such as is seen in fern-like images that look "roughly" the same at small and at large scales. Fractals are produced from an infinite iteration of a finite set of maps, and this algorithm is perfectly suited to the kind of subdivision which is a cornerstone of the discrete wavelet algorithm. Self-similarity could refer alternately to space, and to time. And further versatility is added, in that flexibility is allowed into the definition of "similar."

The idea of a scientific analysis by subdividing a fixed picture or object into its finer parts is not unique to wavelets. It works best for structures with an inherent self-similarity; this self-similarity can arise from numerical scaling of distances.

3.7 Wavelets and Fractals

There are more subtle non-linear self-similarities. The Julia sets in the complex plane (see (2)) are a case in point [BY06], [Bra06], [DL06], [DRS07], [Mil04], [PZ04]. The simplest Julia sets come from a one parameter family of quadratic polynomials $\varphi_c(z) = z^2 + c$, where z is a complex variable and where c is a fixed parameter. The corresponding family of Julia sets J_c have a surprisingly rich structure.

A way to understand them is the following: Consider the two branches of the inverse $\beta_\pm = z \mapsto \pm\sqrt{z - c}$. Then J_c is the unique minimal non-empty compact subset of $\mathbb{C}$, which is invariant under $\{\beta_\pm\}$. (There are alternative ways of presenting J_c but this one fits our purpose. The Julia set J of a holomorphic function, in this case $z \mapsto z^2 + c$, informally consists of those points whose long-time behavior under repeated iteration of substitutions, can change drastically under arbitrarily small perturbations.) Here "long-time" refers to large n, where $\varphi^{(n+1)}(z) = \varphi(\varphi^{(n)}(z))$, $n = 0, 1, ...$, and $\varphi^{(0)}(z) = z$.

Construct Haar wavelets on J_c**.** We are interested in adapting and modifying the Haar wavelet, and the other wavelet algorithms to the Julia sets. The two papers [DJ05], [DJ06b] initiate such a development. In earlier papers [DJ06c], [DJ06a], the authors initiated wavelet transforms for complex fractals. These transforms have some parallels to traditional affine fractals, but subtle non-linearities preclude us from writing down an analogue of Haar wavelets in these different settings. We propose to develop more refined algorithms taking these difficulties into account.

3.8 Fractal Image Processing

Fractal image compression is recently introduced [Bar98] and operates on a set of principles motivated by the theory of iterated function systems (IFS). Unlike wavelets, fractal coders store images as a fixed points of maps on the plane instead of a set of quantized transform coefficients. Fractal compression is related to vector quantization, but fractal coders use a self-referential vector codebook, drawn from the image itself, instead of a fixed codebook. [Dav].

IFS theory motivates a broad class fractal compression schemes but it does not show why particular fractal schemes work well. A wavelet-based framework for ana-

lyzing fractal block coders would simplify the analysis of these codes considerably and give a clear picture of why they are effective. The fractal block coding schemes function essentially by extrapolating Haar wavelet coefficients across scales. This analysis gives insight into various important implementation issues which are vague, namely bit allocation methods, error estimation, quantization vector search strategies, and super-resolution of images.

References

[AK06]   William Arveson and Richard V. Kadison. Diagonals of self-adjoint operators. In *Operator theory, operator algebras, and applications*, volume 414 of *Contemp. Math.*, pages 247–263. Amer. Math. Soc., Providence, RI, 2006.

[APR06]   F. Autin, D. Picard, and V. Rivoirard. Large variance Gaussian priors in Bayesian nonparametric estimation: a maxiset approach. *Math. Methods Statist.*, 15(4):349–373 (2007), 2006.

[APS06]   Sergio Albeverio, Alex Popovici, and Victoria Steblovskaya. A numerical analysis of the extended Black-Scholes model. *Int. J. Theor. Appl. Finance*, 9(1):69–89, 2006.

[Arv07]   William Arveson. Diagonals of normal operators with finite spectrum. *Proc. Natl. Acad. Sci. USA*, 104(4):1152–1158 (electronic), 2007.

[Ash90]   Robert B. Ash. *Information theory*. Dover Publications Inc., New York, 1990. Corrected reprint of the 1965 original.

[Bar96]   Michael F. Barnsley. Fractal image compression. *Notices Amer. Math. Soc.*, 43(6):657–662, 1996.

[Bar98]   Jacquin A. Barnsley, M. F. Application of recurrent iterated function systems to images. *Proc. SPIE*, 1001:122–131, 1998.

[Bar02]   Michael F. Barnsley. Iterated function systems for lossless data compression. In *Fractals in multimedia (Minneapolis, MN, 2001)*, volume 132 of *IMA Vol. Math. Appl.*, pages 33–63. Springer, New York, 2002.

[Bar06]   Michael Fielding Barnsley. *Superfractals*. Cambridge University Press, Cambridge, 2006.

[BB97]   M. Barnsley and L. Barnsley. Fractal image compression. In *Image processing: mathematical methods and applications (Cranfield, 1994)*, volume 61 of *Inst. Math. Appl. Conf. Ser. New Ser.*, pages 183–210. Oxford Univ. Press, Oxford, 1997.

[BEH89]   Michael F. Barnsley, John H. Elton, and Douglas P. Hardin. Recurrent iterated function systems. *Constr. Approx.*, 5(1):3–31, 1989. Fractal approximation.

[Bel90]   Cleary J. G. Witten I. H. Bell, T. C. Prentice Hall, Englewood Cliffs, 1990.

[BHS05]   Michael Barnsley, John Hutchinson, and Örjan Stenflo. A fractal valued random iteration algorithm and fractal hierarchy. *Fractals*, 13(2):111–146, 2005.

[BJ02]   0. Bratelli and P. Jorgensen. *Wavelets Through a Looking Glass: The World of the Spectrum*. Birkhäuser, 2002.

[BKG+]   Paolo Baldi, Kerkyacharian, Gérard, Domenico Marinucci, and Dominique Picard. Subsampling needlet coefficients on the sphere. *arXiv:0706.4169v1 [math.ST]*.

[Bra06]   Mark Braverman. Parabolic Julia sets are polynomial time computable. *Nonlinearity*, 19(6):1383–1401, 2006.

[BY06]   M. Braverman and M. Yampolsky. Non-computable Julia sets. *J. Amer. Math. Soc.*, 19(3):551–578 (electronic), 2006.

[Cas06]   Kutyniok G. Li S. Casazza, P. G. Fusion frames and distributed processing. *preprint*, 2006.

[CFTW06] Peter G. Casazza, Matthew Fickus, Janet C. Tremain, and Eric Weber. The Kadison-Singer problem in mathematics and engineering: a detailed account. In *Operator theory, operator algebras, and applications*, volume 414 of *Contemp. Math.*, pages 299–355. Amer. Math. Soc., Providence, RI, 2006.

[Chr03]   Ole Christensen. *An introduction to frames and Riesz bases*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston Inc., Boston, MA, 2003.

[CK04]   Peter G. Casazza and Gitta Kutyniok. Frames of subspaces. In *Wavelets, frames and operator theory*, volume 345 of *Contemp. Math.*, pages 87–113. Amer. Math. Soc., Providence, RI, 2004.

[Con94]   Alain Connes. *Noncommutative geometry*. Academic Press Inc., San Diego, CA, 1994.

[Cun77]   Joachim Cuntz. Simple C^*-algebras generated by isometries. *Comm. Math. Phys.*, 57(2):173–185, 1977.

[CW84]   John G. Cleary and Ian H. Witten. A comparison of enumerative and adaptive codes. *IEEE Trans. Inform. Theory*, 30(2, part 2):306–315, 1984.

[D'A97]   Sacchi M. F. D'Ariano, G. M. Optical von neumann measurement. *Phys. Lett. A*, 231(5-6):325–330, 1997.

[Dau92]   Ingrid Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. 1992.

[Dau93]   Ingrid Daubechies. *Wavelet transforms and orthonormal wavelet bases*. Proc. Sympos. Appl. Math. 1993.

[Dav]   G. Davis. *Why Fractal Block Coders Work*, volume 159 of *NATO ASI Series Computer and Systems Sciences*.

[DJ05]   Dorin Ervin Dutkay and Palle E. T. Jorgensen. Wavelet constructions in non-linear dynamics. *Electron. Res. Announc. Amer. Math. Soc.*, 2005.

[DJ06a]   Dorin E. Dutkay and Palle E. T. Jorgensen. Wavelets on fractals. *Rev. Mat. Iberoamericana*, 22, 2006.

[DJ06b]   Dorin Ervin Dutkay and Palle E. T. Jorgensen. Hilbert spaces built on a similarity and on dynamical renormalization. *J. Math. Phys.*, 2006.

[DJ06c]   Dorin Ervin Dutkay and Palle E. T. Jorgensen. Iterated function systems, Ruelle operators, and invariant projective measures. *Math. Comp.*, 2006.

[DL92]   Ingrid Daubechies and Jeffrey C. Lagarias. Two-scale difference equations. II. Local regularity, infinite products of matrices and fractals. *SIAM J. Math. Anal.*, 1992.

[DL98]   Xingde Dai and David R. Larson. Wandering vectors for unitary systems and orthogonal wavelets. *Mem. Amer. Math. Soc.*, 134(640):viii+68, 1998.

[DL06]   Robert L. Devaney and Daniel M. Look. A criterion for Sierpinski curve Julia sets. *Topology Proc.*, 30(1):163–179, 2006. Spring Topology and Dynamical Systems Conference.

[DLS98]   Xingde Dai, David R. Larson, and Darrin M. Speegle. Wavelet sets in $\mathbf{R}^n$. II. In *Wavelets, multiwavelets, and their applications (San Diego, CA, 1997)*, volume 216 of *Contemp. Math.*, pages 15–40. Amer. Math. Soc., Providence, RI, 1998.

[DRS07]   Robert L. Devaney, Mónica Moreno Rocha, and Stefan Siegmund. Rational maps with generalized Sierpinski gasket Julia sets. *Topology Appl.*, 154(1):11–27, 2007.

[DS06]   S. Dereich and M. Scheutzow. High-resolution quantization and entropy coding for fractional Brownian motion. *Electron. J. Probab.*, 11:no. 28, 700–722 (electronic), 2006.

[DVDD98] David L. Donoho, Martin Vetterli, R. A. DeVore, and Ingrid Daubechies. Data compression and harmonic analysis. *IEEE Trans. Inform. Theory*, 44(6):2435–2476, 1998. Information theory: 1948–1998.

[Eld03]   Yonina C. Eldar. von Neumann measurement is optimal for detecting linearly independent mixed quantum states. *Phys. Rev. A (3)*, 68(5):052303, 4, 2003.

[GWE04]   R. C. Gonzalez, R. E. Woods, and S. L. Eddins. *Digital Image Processing Using MATLAB*. Prentice Hall, 2004.

[HK06]   Kais Hamza and Fima C. Klebaner. On nonexistence of non-constant volatility in the Black-Scholes formula. *Discrete Contin. Dyn. Syst. Ser. B*, 6(4):829–834 (electronic), 2006.

[HL00]   Deguang Han and David R. Larson. Frames, bases and group representations. *Mem. Amer. Math. Soc.*, 147(697):x+94, 2000.

[JMR01]   Stéphane Jaffard, Yves Meyer, and Robert D. Ryan. *Wavelets*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, revised edition, 2001. Tools for science & technology.

[Jor06]   Palle E. T. Jorgensen. *Analysis and probability: wavelets, signals, fractals*, volume 234 of *Graduate Texts in Mathematics*. Springer, New York, 2006.

[JP98a]   Palle E. T. Jorgensen and Steen Pedersen. Dense analytic subspaces in fractal L^2-spaces. *J. Anal. Math.*, 75:185–228, 1998.

[JP98b]   Palle E. T. Jorgensen and Steen Pedersen. Local harmonic analysis for domains in $\mathbf{R}^n$ of finite measure. In *Analysis and topology*, pages 377–410. World Sci. Publ., River Edge, NJ, 1998.

[JP98c]   Palle E. T. Jorgensen and Steen Pedersen. Orthogonal harmonic analysis and scaling of fractal measures. *C. R. Acad. Sci. Paris Sér. I Math.*, 326(3):301–306, 1998.

[JS07a]   P. E. T. Jorgensen and M. S. Song. Comparison of discrete and continuous wavelet transforms. *Springer Encyclopedia of Complexity and Systems Science*, 2007.

[JS07b]   Palle E. T. Jorgensen and Myung-Sin Song. Entropy encoding, hilbert space, and karhunen-loève transforms. *J. Math. Phys.*, 48(10):103503, 2007.

[JT06]   Svante Janson and Johan Tysk. Feynman-Kac formulas for Black-Scholes-type operators. *Bull. London Math. Soc.*, 38(2):269–282, 2006.

[Jum05]   Guy Jumarie. Merton's model of optimal portfolio in a Black-Scholes market driven by a fractional Brownian motion with short-range dependence. *Insurance Math. Econom.*, 37(3):585–598, 2005.

[Kig01]   Jun Kigami. *Analysis on fractals*, volume 143 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2001.

[Kig03]   Jun Kigami. Harmonic analysis for resistance forms. *J. Funct. Anal.*, 204(2):399–444, 2003.

[KP04]   Gérard Kerkyacharian and Dominique Picard. Regression in random design and warped wavelets. *Bernoulli*, 10(6):1053–1105, 2004.

[KS59]   Richard V. Kadison and I. M. Singer. Extensions of pure states. *Amer. J. Math.*, 81:383–400, 1959.

[Lar07]   David R. Larson. Unitary systems and wavelet sets. In *Wavelet analysis and applications*, Appl. Numer. Harmon. Anal., pages 143–171. Birkhäuser, Basel, 2007.

[LP06]   Michel L. Lapidus and Erin P. J. Pearse. A tube formula for the Koch snowflake curve, with applications to complex dimensions. *J. London Math. Soc. (2)*, 74(2):397–414, 2006.

[LSST06]   David Larson, Eckart Schulz, Darrin Speegle, and Keith F. Taylor. Explicit cross-sections of singly generated group actions. In *Harmonic analysis and applications*, Appl. Numer. Harmon. Anal., pages 209–230. Birkhäuser Boston, Boston, MA, 2006.

[LTW06]   David R. Larson, Wai-Shing Tang, and Eric Weber. Robertson-type theorems for countable groups of unitary operators. In *Operator theory, operator algebras, and applications*, volume 414 of *Contemp. Math.*, pages 289–295. Amer. Math. Soc., Providence, RI, 2006.

[Mey06]   Gunter H. Meyer. The Black Scholes Barenblatt equation for options with uncertain volatility and its application to static hedging. *Int. J. Theor. Appl. Finance*, 9(5):673–703, 2006.

[Mil04]   John Milnor. Pasting together Julia sets: a worked out example of mating. *Experiment. Math.*, 13(1):55–92, 2004.

[MST99]   Yves Meyer, Fabrice Sellan, and Murad S. Taqqu. Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. *J. Fourier Anal. Appl.*, 5(5):465–494, 1999.

[OS07]   Kasso A. Okoudjou and Robert S. Strichartz. Asymptotics of eigenvalue clusters for Schrödinger operators on the Sierpiński gasket. *Proc. Amer. Math. Soc.*, 135(8):2453–2459 (electronic), 2007.

[Pet86]   Karl Petersen. Chains, entropy, coding. *Ergodic Theory Dynam. Systems*, 6(3):415–448, 1986.

[PZ04]   C. L. Petersen and S. Zakeri. On the Julia set of a typical quadratic polynomial with a Siegel disk. *Ann. of Math. (2)*, 159(1):1–52, 2004.

[RM06]   Marianito R. Rodrigo and Rogemar S. Mamon. An alternative approach to solving the Black-Scholes equation with time-varying parameters. *Appl. Math. Lett.*, 19(4):398–402, 2006.

[SCE01]   A. Skodras, C. Christopoulos, and T. Ebrahimi. Jpeg 2000 still image compression standard. *IEEE Signal Processing Magazine*, 18:36–58, Sept. 2001.

[Son07]   M. S. Song. Entropy encoding in wavelet image compression. *Representations, Wavelets and Frames A Celebration of the Mathematical Work of Lawrence Baggett*, pages 293–311, 2007.

[Str06a]   Robert S. Strichartz. Convergence of mock Fourier series. *J. Anal. Math.*, 99:333–353, 2006.

[Str06b]   Robert S. Strichartz. *Differential equations on fractals*. Princeton University Press, Princeton, NJ, 2006. A tutorial.

[Use01]   B. E. Usevitch. A tutorial on modern lossy wavelet image compression: Foundations of jpeg 2000. *IEEE Signal Processing Magazine*, 18:22–35, Sept. 2001.

[Wal99]   J. S Walker. *A Primer on Wavelets and their Scientific Applications*. Chapman & Hall, CRC, 1999.

[Wan04]   M. S. Wang. Classical limit of von Neumann measurement. *Phys. Rev. A (3)*, 69(3):034101, 2, 2004.

[Wit04]   I. H. Witten. Adaptive text mining: inferring structure from sequences. *J. Discrete Algorithms*, 2(2):137–159, 2004.

[WNC87] I. H. Witten, R. M. Neal, and J. G. Cleary. Arithmetic coding for data compression. *Comm. of the ACM*, 30(6):520–540, 1987.