FROM FRACTAL IMAGE COMPRESSION TO FRACTAL-BASED METHODS IN MATHEMATICS

EDWARD R. VRSCAY*

1. Introduction. In keeping with the philosophy of this workshop, the aim of this presentation is to provide an overview of the research done over the years at Waterloo on fractal-based methods of approximation and associated inverse problems. Near the end, some new and encouraging results regarding "fractal enhancement" will be presented. The paper concludes with thoughts and challenges on how the mathematical methods that underlie fractal image compression could be used in other areas of mathematics.

Let us go back to first principles for a moment in order to recall some of the early thinking behind fractal image compression (FIC). In fact, since the early work of Barnsley, Jacquin *et al.*, there has been very little change in the basic idea of FIC. Most efforts have focussed on developing strategies to perform "collage coding" as effectively as possible – whether it be in the pixel or wavelet domain. This includes the the competition between employing the largest possible domain pools and searching them as quickly as possible.

Very briefly, let I denote an image of interest as defined by an *image function* $u(x,y)$ supported over a region $X \in \mathbf{R}^2$ as shown in Figure 1 below. Here $x, y \in X$ denote spatial coordinates of a point or pixel of I.

Now suppose that there exists a suitable partition of D into subblocks R_i so that $X = \cup_i R_i$. For simplicity, the R_i are assumed to be "nonoverlapping," intersecting only at their boundaries, which are assumed to have zero Lebesgue measure in the plane. (In the discrete case, i.e., pixels, there is no overlapping.) Assume that associated with each subblock R_i is a larger subblock $D_i \subseteq X$ so that $R_i = w_i(D_i)$, where w_i is a 1-1 contraction map. Furthermore, we have found that the image function $u(R_i)$ supported on R_i is well approximated by a modified copy of the image function $u(D_i)$ as follows:

$$u(R_i) \cong \phi_i(u(D_i)) = \phi_i(u(w_i^{-1}(R_i))), \tag{1}$$

where $\phi_i : \mathbf{R} \to \mathbf{R}$ is a *greyscale* or *grey level* map that operates on the pixel intensities. The process is illustrated in Figure 2 below.

Because of the nonoverlapping nature of the partition, we may write

$$u(x,y) \cong (Tu)(x,y) = \sum_i \phi_i(u(w_i^{-1}(x,y))). \tag{2}$$

*Department of Applied Mathematics, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1 (ervrscay@links.uwaterloo.ca); http://links.uwaterloo.ca

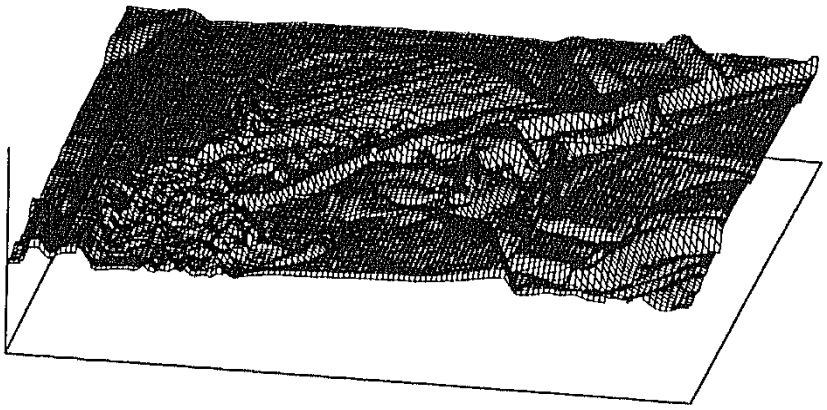

FIG. 1. *Image function $u(x,y)$ associated with 8-bit (256 level) Lena image.*

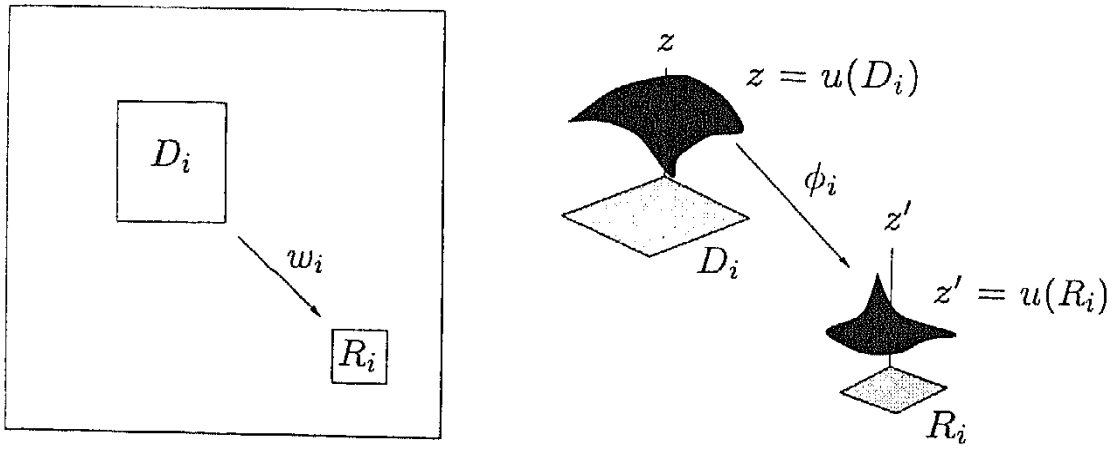

FIG. 2. *Illustration of the fractal transform. Left: Range block R_i and associated domain block D_i. Right: Greyscale mapping ϕ_i from $u(D_i)$ to $u(R_i)$.*

In other words, the image function u is approximated by a union of spatially-contracted (w_i) and greyscale-distorted (ϕ_i) copies of itself. We may consider this union of modified copies as defining a special kind of *fractal transform* operator T. If the above approximation is a good one, then the so-called *collage distance* $\| u - Tu \|$ is small.

The early excitement of fractal image compression, following the results of IFS theory, lay in the fact that the operator T, under suitable conditions, can be made contractive. This implies the existence of a unique fixed point function $\bar{u} = T\bar{u}$ that can be generated by the iteration process $u_{n+1} = Tu_n$, with u_0 arbitrary, e.g. a blank computer screen. Moreover, a simple triangle inequality – the so-called "Collage theorem" –establishes that if u is "close" to Tu, then u is also "close" to $\bar{u}$, implying that $\bar{u}$ is a good approximation to u. The dramatic conclusion was that one needs only to store the parameters that define T in order to generate the approximation $\bar{u}$.

From this early work arose much intense effort in fractal image compression. Indeed, the points outlined above provided a recurring theme for research at Waterloo which can be summarized as follows:

1. The identificaton of suitable metric spaces (Y, d_Y) of functions, measures or distributions that can be used to represent "images" supported on a region X.
2. The construction of suitable fractal transform operators $T : Y \to Y$ over these spaces. Such operators produce spatially contracted and greyscale distorted copies of an image function $u \in Y$ – the *fractal components* $\phi_i(u(w_i^{-1}(R_i)))$ in Eqs. (1) and (2) – and then recombine these components in an appropriate manner, in particular when more than one fractal component may exist at a point $x \in X$. It is also desirable that, under certain conditions on the w_i and the ϕ_i, the operator T is contractive in (Y, d_Y).
3. The formulation and solution of inverse problems in these spaces, where a "target" element $u \in Y$ is approximated by a fixed point $\bar{u}$ of a contractive fractal transform T.
4. The construction of fractal transform operators on other metric spaces that may not necessarily be connected with image function spaces.

Finally, we consider the more general question whether the mathematical methods behind inverse problems in (3) above can be used in other, possibly "non-fractal", areas of mathematics. Some encouraging results that indicate an affirmative answer will be presented.

1.1. Iterated function systems and IFS with probabilities. Let us now backtrack even further to summarize the most important features of Iterated Function Systems. Such a review obviously establishes notation and refreshes the reader's memory of IFS basics. Here, however, the review will also allow us to appreciate the recurring themes that make their way into later IFS research and its applications. Note that it is not possible here to provide a complete account of works that could be considered as precursors to IFS theory. At the risk of omitting some important works, I simply mention a few references, beginning with the work of Williams [67] who considered the problem of fixed points of finite compositions of contractive maps. Nadler [55] considered systems of maps as defining "multifunctions" following earlier work of others. In his studies of learning models, Karlin [39] constructed random walks over Cantor-like sets, essentially the "Chaos Game" [5] of IFS.

Our review begins with the work of Hutchinson [36] who studied the geometric and measure theoretic aspects of systems of contractive maps with associated probabilities, and the use of such systems to construct fractal sets. Barnsley and Demko [6] independently discovered the use of such systems of mappings, which they referred to as "Iterated Function Systems", but in a more probabilistic setting.

Let (X,d) be a complete metric space (typically $[0,1]^n$ with usual Euclidean metric) and let $\mathcal{H}(X)$ denote the set of non-empty subsets of X. The metric space $(\mathcal{H}(X),h)$, where h denotes the Hausdorff metric, is complete [20]. Now let $w_i : X \to X$, $1 \le i \le N$, denote a set of contraction maps on X with contraction factors $c_i \in [0,1)$ defined in the usual way:

$$c_i := \sup_{x,y \in X, x \neq y} d(w_i(x), w_i(y))/d(x,y). \tag{3}$$

Associated with each contraction map w_i is a set-valued mapping $\hat{w}_i : \mathcal{H}(X) \to \mathcal{H}(X)$ defined by $\hat{w}_i(S) = \{w_i(x) | x \in S\}$ for $S \in \mathcal{H}(X)$. The IFS operator $\hat{\mathbf{w}}$ associated with the N-map IFS $\mathbf{w}$ is defined as follows:

$$\hat{\mathbf{w}}(S) = \bigcup_{i=1} \hat{w}_i(S), \quad S \in \mathcal{H}(X). \tag{4}$$

This operator $\hat{\mathbf{w}}$ is contractive on $(\mathcal{H}(X),h)$ [36]:

$$h(\hat{\mathbf{w}}(A), \hat{\mathbf{w}}(B) \le ch(A,B), \quad \forall A, B \in \mathcal{H}(X), \tag{5}$$

where $c = \max_{1 \le i \le N}\{c_i\} < 1$. The completeness of $(\mathcal{H}(X),h)$ guarantees the existence of a unique fixed point $A = \hat{\mathbf{w}}(A)$ in $\mathcal{H}(X)$. The set A, also called the *attractor*, can be viewed as the IFS representation of an image. From Eq. (4), it satisfies the self-tiling property,

$$A = \bigcup_{i=1} \hat{w}_i(A). \tag{6}$$

The union is a natural operator to combine the component sets $A_i = \hat{w}_i(A)$. (Note that this definition was later extended to "partitioned" or "local" IFS, where the contractions w_i could map *subsets* $D_i \subseteq X$ to smaller subsets $R_i \subset X$. We omit the details.)

The moral of the story: "Classical" IFS generate *sets*. Sets can be interpreted as images but only of a limited class – black/white images. A more useful method would accomodate *shading*, i.e., a greater number of greyscale values in the interval [0,1] (or [0,255]). Barnsley and coworkers, beginning in [6], realized that invariant measures of IFS with probabilities could perform this task.

Let $\mathbf{p} = \{p_1, p_2, \ldots, p_N\}$ denote a set of probabilities associated with the IFS maps w_i, with $\sum_i p_i = 1$. Let $\mathcal{M}(X)$ denote the space of probability measures on the Borel sigma field of X with the following Monge-Kantorovich metric [36]

$$d_M(\mu,\nu) = \sup_{f \in Lip_1(X)} \left| \int f d\mu - \int f d\nu \right|, \tag{7}$$

where $Lip_1(X)$ denotes the set of Lipschitz functions $f : X \to \mathbf{R}$ with Lipschitz constant $K \le 1$. Associated with this *IFS with probabilities*

(IFSP) $(\mathbf{w}, \mathbf{p})$ is an operator $M : \mathcal{M}(X) \to \mathcal{M}(X)$ defined as follows. For a $\mu \in \mathcal{M}(X)$,

$$(M\mu)(S) = \sum_i p_i \mu(w_i^{-1}(S)), \qquad \forall S \in \mathcal{H}(X). \tag{8}$$

Then [36]

$$d_M(M\mu, M\nu) \leq c d_M(\mu, \nu), \qquad \forall \mu, \nu \in \mathcal{M}(X). \tag{9}$$

Since M is a contraction mapping in $\mathcal{M}(X)$, there exists a unique measure $\bar{\mu} \in \mathcal{M}(X)$ such that $\bar{\mu} = M\bar{\mu}$. This *invariant measure* of the IFSP $(\mathbf{w}, \mathbf{p})$ also satisfies a self-tiling property:

$$\bar{\mu}(S) = \sum_i p_i \bar{\mu}(w_i^{-1}(S)), \qquad \forall S \in \mathcal{H}(X). \tag{10}$$

With IFSP, images are represented as probability measures supported on the attractor $A \subseteq X$. The measure $\mu(B)$ of an appropriate Borel subset $B \subseteq X$ can be interpreted as the grey level value of the pixel representing B. The probabilities p_i can be used to modify these values, hence rendering the *shading* of the attractor A.

Most probably the reader automatically accepted the use of the summation function in Eq. (8): If a point $x \in X$ has several preimages, say $w_{i_k}^{-1}(x), 1 \leq k \leq n$, then we linearly combine the measures $\mu(w_{i_k}^{-1}(x))$ to obtain $(M\mu)(x)$, using the probabilities p_i as weights. This procedure follows naturally from both probabilistic as well as measure theoretic aspects.

Jacquin's original "block IFS" method [37] was formulated in the language of measure preserving transformations and invariant measures, quite reminiscent of earlier IFSP work. In its applications to image compression, the method operates on measures over discrete pixel domains – scaling them and adding them, which is essentially equivalent to treating the images as functions. Indeed, it was quite natural that later work on fractal image compression, for example that of Fisher [22], considered images as functions – in particular, $\mathcal{L}^2$ functions, the usual procedure in signal and image processing. The space $\mathcal{L}^\infty$ offers some attractive simplifications for fractal image compression. For example, the contractivity of the fractal transform T depends only upon the Lipschitz factors of the grey level maps ϕ_i, unlike the case for $\mathcal{L}^2$ [9].

1.2. IFS on $\mathcal{L}^p$ spaces. An IFS-type method can easily be formulated over the space $\mathcal{L}^p(X), p \geq 1$, of Lebesgue integrable functions on X:

$$\mathcal{L}^p(X) = \{f : X \to \mathbf{R} \mid \int_X |f(x)|^p dx < \infty\}. \tag{11}$$

As before, let w_i, $1 \leq i \leq N$, denote a set of IFS contraction maps on X. Associated with each map w_i is a *grey level* map $\phi_i : \mathbf{R} \to \mathbf{R}$ satisfying a Lipschitz property on $\mathbf{R}$: There exists a $K_i \geq 0$ such that

$$|\phi_i(t_1) - \phi_i(t_2)| \leq K_i|t_1 - t_2|, \quad \forall\, t_1, t_2 \in \mathbf{R}. \tag{12}$$

In most applications, the grey level maps ϕ_i are affine, i.e., $\phi_i(t) = \alpha_i t + \beta_i$, so that $K_i = |\alpha_i|$.

The *IFS with grey level maps* (IFSM) $(\mathbf{w}, \Phi)$ defines an operator $T : \mathcal{L}^p(X) \to \mathcal{L}^p(X)$ as follows: For $u \in \mathcal{L}^p(X)$,

$$(Tu)(x) = \sum_i \phi_i(u(w_i^{-1}(x))). \tag{13}$$

A simple calculation [27] shows that for all $u, v \in \mathcal{L}^p(X)$,

$$\| Tu - Tv \|_p \leq C_p \| u - v \|_p, \quad C_p = \sum_i c_i^{1/p} K_i. \tag{14}$$

If $C_p < 1$, then T is contractive on $\mathcal{L}^p(X)$. The unique fixed point function $\bar{u} = T\bar{u}$ satisfies the functional equation

$$\bar{u}(x) = \sum_i \phi_i(\bar{u}(w_i^{-1}(x))), \tag{15}$$

a self-similarity relation in which $\bar{u}$ is expressed as a union of contracted (in the x direction) and distorted (in the grey level direction) copies of itself.

In the special case that the sets $X_i = \hat{w}_i(X)$ are *nonoverlapping* (or overlapping on a set of Lebesgue measure 0), then

$$\| Tu - Tv \|_p \leq \bar{C}_p \| u - v \|_p, \quad \bar{C}_p = [\sum_i c_i K_i^p]^{1/p}. \tag{16}$$

Note that

$$\bar{C}_p \leq C_p \leq K, \quad K = \max_i\{K_i\}. \tag{17}$$

In the nonoverlapping case, with $p = \infty$,

$$\| Tu - Tv \|_\infty \leq K \| u - v \|_\infty, \quad u, v \in \mathcal{L}^\infty(X). \tag{18}$$

This is the usual bound presented in the literature on fractal transforms [9].

In many treatments, both the IFS maps w_i and the grey level maps ϕ_i are assumed to be affine. In such cases, we refer to the IFSM $(\mathbf{w}, \Phi)$ as an *affine IFSM*. In the case that X is one-dimensional, e.g. $X = [0, 1]$, the maps will have the form

$$w_i = s_i x + a_i, \qquad \phi_i(t) = \alpha_i t + \beta_i, \quad 1 \leq i \leq N, \tag{19}$$

where $\alpha_i, \beta_i \in \mathbf{R}$ and $|s_i| < 1$.

Another interesting form of an IFSM operator is the following:

$$(Tu)(x) = \theta(x) + \sum_i \alpha_i u(w_i^{-1}(x)). \tag{20}$$

Here, the function $\theta(x)$ acts as a "condensation" function for the IFSM. Note that the affine IFSM of Eq. (19) is a special case of IFSM with condensation, with

$$\theta(x) = \sum_i \beta_i I_{w_i(X)}(x), \tag{21}$$

where $I_S(x)$ denotes the characteristic function of a set $S \subseteq X$.

In closing we mention that the form of the grey level maps may be generalized to include *place-dependent* maps $\phi_i : \mathbf{R} \times X \to \mathbf{R}$, where the ϕ_i are dependent upon both the grey level value at a preimage as well as the spatial location of the preimage itself. (This is analogous to IFS with place-dependent probabilities [7].) Much of the above theory for IFSM extends easily to place-dependent IFSM, as shown in [28]. This is the basis of the "Bath Fractal Transform," a very effective method of coding images [53, 54, 68].

1.3. Self-similar functions. Finally, a note on Eq. (15) satisfied by the IFSM fixed point attractor function $\bar{u}$. The notion of self-similarity properties of functions is certainly not foreign in mathematics, especially in the field of wavelet theory. Let $\phi(x)$ be a *scaling function* (not to be confused with the ϕ_i notation used for grey level functions) that satisfies the conditions for a multiresolution analysis [47]. Now define the space $V_0 = \text{span}\{\phi(x-n)\}_{n=-\infty}^{\infty}$. Then $V_0 \subset V_1 = \text{span}\{\phi(2x-n)\}_{n=\infty}^{\infty}$, implying that ϕ satifies the functional equation

$$\phi(x) = \sum_k c_k \phi(2x - k). \tag{22}$$

The functions $\phi(2x - k)$ are dilations of $\phi(x)$ produced by the IFS-type contraction maps $w_k(x) = \frac{1}{2}(x+k)$. One can view the scaling function $\phi(x)$ as the fixed point of an IFSM operator with grey level maps $\phi_k(t) = c_k t$. (In this case, the IFSM operator is linear, implying that $\phi(x) = 0$ is also a fixed point solution.)

A standard definition of self-similarity [38], which we include here for completeness, demands that a function f be self-similar over sets satisfying an "open set condition" used in the case for measures. Suppose that f is continuous and compactly supported and let Ω be the bounded open subset of $\mathbf{R}$ such that $\overline{\Omega} = \text{supp}\{f\}$. Further suppose that there exist disjoint subsets $\Omega_i \subset \Omega$, $i = 1, 2, \ldots, N$, with $S_i(\Omega) = \Omega_i$, where the S_i are contractive similitudes. Then

$$f(x) = \sum_i \lambda_i f(S_i^{-1}(x)) + g(x), \tag{23}$$

where $g(x) = g_j(x)$ is Lipschitz for $x \in \Omega_j$ and $g(x) = f(x)$ for $x \notin \cup_i \Omega_i$. From Eq. (20), f is the fixed point of an N-map (nonoverlapping) IFSM with condensation function g.

The idea of self-similarity can be traced farther back in history. For example, Bajraktarevic [2], following the work of Read [56], studied a class of functional equations of the form

$$f(x) = F\big(x, f(g_1(x)), \ldots, f(g_N(x))\big). \tag{24}$$

This class includes the functional equations satisfied by fixed points of IFS-type operators used in fractal image coding. For this reason, various authors have referred to fractal transforms as Read-Bajraktarevic operators. The various forms of fractal interpolation functions, for example in [4, 18], demonstrate self-similarity. Indeed, the construction of orthogonal wavelets from fractal interpolation functions has also received much attention – see [49] and references therein. A generalized self-similarity relation of $\mathcal{L}^p$ functions relevant to the construction of wavelets and multiwavelets has also recently been discussed [12].

2. Unification of IFS methods on function/measure spaces. In this section we summarize the formulation of IFS methods on various function spaces [27]. This leads to the definition of fractal transforms on the space of distributions which permits a unification of IFS methods on function and probability measure spaces [28].

Early work sought to complement the work of Fisher [22, 23], who formulated fractal transforms on $\mathcal{L}^2$ functions. Our first attempts were to look for function spaces that would appear to be well suited to IFS geometric contractions. Since the IFS was formulated in terms of the Hausdorff metric, we were inspired to consider the space of *fuzzy sets*, where the distance between two functions involves the Hausdorff distances between their respective level sets.

2.1. Fuzzy sets and IFZS. The space of fuzzy sets on X, $\mathcal{F}^*(X)$, is defined to be the set of all functions $u : X \to [0,1]$ that are (1) upper semicontinuous on (X,d) and (2) normalized, i.e., for each $u \in \mathcal{F}^*(X)$ there exists at least one point $x_0 \in X$ for which $u(x_0) = 1$. The grey level range is $R_g = [0,1]$. The metric on this space is given by

$$d_\infty(u,v) = \sup_{0\le\alpha\le 1} h([u]^\alpha, [v]^\alpha), \quad \forall u, v \in \mathcal{F}^*(X), \tag{25}$$

where the α-level sets of $u \in \mathcal{F}^*(X)$ for $\alpha \in [0,1]$ are defined as follows:

$$\begin{aligned} [u]^\alpha &:= \{x \in X : u(x) \ge \alpha\}, \quad \alpha \in (0,1], \\ [u]^0 &:= \text{closure}\{x \in X : u(x) > 0\}. \end{aligned} \tag{26}$$

The metric space $(\mathcal{F}^*(X), d_\infty)$ is complete [17].

An IFS-type method on $\mathcal{F}^*(X)$, referred to as IFZS [11], may be formulated as follows. Associated with the IFS contraction maps $\mathbf{w} = \{w_1, \ldots, w_N\}$ is a set of grey level maps $\Phi = \{\phi_1, \phi_2, ..., \phi_N\}$, $\phi_i : [0,1] \to$

[0,1], such that for all $i \in \{1,2,\ldots,N\}$: (i) ϕ_i is nondecreasing on [0,1], (ii) ϕ_i is right continuous on [0,1), (iii) $\phi_i(0) = 0$ and (iv) $\phi_{i^*}(1) = 1$ for at least one $i^* \in \{1,2,...,N\}$.

The pair of vectors $(\mathbf{w}, \Phi)$ comprise an IFZS with associated fractal transform operator $T : \mathcal{F}^*(X) \to \mathcal{F}^*(X)$. For a $u \in \mathcal{F}^*(X)$, its image Tu is (up to a minor technicality which will be ignored here, cf. [11]) given by

$$(Tu)(x) = \sup_{1 \le i \le N} \{\phi_i(u(w_i^{-1}(x)))\}, \quad \forall x \in X. \tag{27}$$

The contractivity of the IFS maps w_i implies that T is a contraction map on $(\mathcal{F}^*(X), d_\infty)$ since [11]

$$d_\infty(Tu, Tv) \le c d_\infty(u, v), \quad \forall u, v \in \mathcal{F}^*(X). \tag{28}$$

The completeness of this space guarantees the existence of a unique fixed point $\bar{u} \in \mathcal{F}^*(X)$. From the definition of T in Eq. (27), the α-level sets of $\bar{u}$ obey the following generalized self-tiling property:

$$[\bar{u}]^\alpha = \bigcup_{i=1} w_i([\phi_i \circ \bar{u}]^\alpha), \quad \alpha \in [0,1]. \tag{29}$$

2.2. From IFS to IFZS. For completeness, it should be noted that "normal" IFS on sets can be viewed as a special case of IFZS. The connection between sets and functions is accomplished by means of characteristic functions [27]. Let $I_S(x)$ denote the characteristic function of a set $S \in \mathcal{H}(X)$, i.e., $I_S(x) = 1$ if $x \in S$ and 0 otherwise. Now let $A, B \in \mathcal{H}(X)$ and $C = A \cup B \in \mathcal{H}(X)$. Then

$$I_C(x) = \sup\{I_A(x), I_B(x)\}. \tag{30}$$

From the property that $I_{\tilde{w}_i(S)}(x) = I_S(w_i^{-1}(x))$, it follows that

$$I_{\tilde{\mathbf{w}}(S)}(x) = \sup_{1 \le i \le N} \{I_S(w_i^{-1}(x))\}. \tag{31}$$

It is natural to consider the normalized ("black and white") function space $\mathcal{F}^*_{BW}(X) \subset \mathcal{F}^*(X)$ defined as

$$\mathcal{F}^*_{BW}(X) = \{u : X \to \{0,1\} \mid \mathrm{supp}(u) \in \mathcal{H}(X)\}. \tag{32}$$

This leads to the following fractal transform operator $T : \mathcal{F}^*_{BW}(X) \to \mathcal{F}^*_{BW}(X)$ associated with an N-map IFS $\mathbf{w}$:

$$(Tu)(x) = \sup_{1 \le i \le N} \{u(w_i^{-1}(x))\}, \quad \forall x \in X, \tag{33}$$

a special case of the IFZS method in which the grey level maps are identity maps.

2.3. From IFZS to IFSM. Unfortunately, the IFZS method is difficult to work with. The grey level maps must satisfy some rather restrictive conditions that lead to some interesting theoretical consequences [24]. The Hausdorff metric d_∞ is also very restrictive from both practical (image processing) as well as theoretical perspectives – see [27] for a discussion. In [27], the idea was to continue to consider functions in terms of their α-level sets as these sets are transformed by geometric (w_i) and grey level (ϕ_i) maps. However, a weaker measure of distance between these sets, namely the symmetric difference, was examined. Briefly:

1. For a $\mu \in \mathcal{M}(X)$ and $u, v \in \mathcal{F}^*(X)$, define for all $\alpha \in [0,1]$,

$$\begin{aligned} G(u,v;\alpha) &= \mu([u]^\alpha \triangle [v]^\alpha) \\ &= \int_X |I_{[u]^\alpha}(x) - I_{[v]^\alpha}(x)| d\mu(x), \end{aligned} \tag{34}$$

   where $\triangle$ denotes the symmetric difference operator: For $A, B \subseteq X$, $A \triangle B = (A \cup B) - (A \cap B)$.
2. Now let ν be a finite measure on $\mathcal{B}(R_g)$, the Borel sets on a suitable *grey level range* $R_g \subset \mathbf{R}$, and define

$$g(u,v;\nu) = \int_{R_g} G(u,v;\alpha) d\nu(\alpha). \tag{35}$$

   An application of Fubini's theorem yields

$$\begin{aligned} g(u,v;\nu) &= \nu(\{0\})\mu([u]^0 \triangle [v]^0) + \int_{X_u} \nu((v(x),u(x)])d\mu(x) \\ &\quad + \int_{X_v} \nu((u(x),v(x)])d\mu(x), \end{aligned} \tag{36}$$

   where $X_u = \{x \in X \mid u(x) > v(x)\}$ and $X_v = \{x \in X \mid v(x) > u(x)\}$.

It can be shown that $g(u,v;\nu)$ is a pseudometric on $\mathcal{L}^1(X,\mu)$. In the particular case that $\nu = m$, the Lebesgue measure on the grey level range R_g, Eq. (36) becomes

$$g(u,v;m) = \int_X |u(x) - v(x)| d\mu(x), \tag{37}$$

the $\mathcal{L}^1(X,\mu)$ distance between u and v. The restrictive Hausdorff metric d_∞ over α-level sets has been replaced by a weaker pseudometric (metric on the measure algebra) involving integrations over X and R_g. The result is a fractal transform method on the function space $\mathcal{L}^1(X,\mu)$. While it appears that only the $\mathcal{L}^1$ distance can be generated by a measure ν on $\mathcal{B}(R_g)$, it is still natural to consider fractal transforms over the general function spaces $\mathcal{L}^p(X,\mu)$.

The result of the above is a union of IFS methods for various image function spaces. What now remains is to establish a connection with IFSP on probability measures. This is accomplished by means of an IFS-method over a space of distributions that includes both integrable functions and measures. This is summarized in the next section.

In closing, we mention that no subsequent work has been done to explore the possible use of the the distance function g in Eq. (36) as an image metric – in particular, the role of the greyscale measure ν.

2.4. IFS on the space of distributions $\mathcal{D}'(X)$. In what follows $X = [0,1]$ although the extension to $[0,1]^n$ is straightforward. Distributions [61] are defined as linear functionals over a suitable space of "test functions", to be denoted as $\mathcal{D}(X)$. Here we choose $\mathcal{D}(X) = C^\infty(X)$, the space of infinitely differentiable real-valued functions on X. (Note: In the literature, $\mathcal{D}(X)$ is normally taken to be $C_0^\infty(X)$, the set of $C^\infty(X)$ functions with compact support on X. With this choice, the expressions for distributional derivatives simplify due to the vanishing of boundary terms.) The space of distributions on X, to be denoted as $\mathcal{D}'(X)$, is the set of all bounded linear functionals on $\mathcal{D}(X)$, that is, $F : \mathcal{D}(X) \to \mathbf{R}$, such that

1. $|F(\psi)| < \infty$ for all $\psi \in \mathcal{D}(X)$,
2. $F(c_1\psi_1 + c_2\psi_2) = c_1F(\psi_1) + c_2F(\psi_2)$, $c_1, c_2 \in \mathbf{R}$, $\psi_1, \psi_2 \in \mathcal{D}(X)$.

The space $\mathcal{D}'(X)$ will accomodate the $\mathcal{L}^p(X)$ spaces $1 \le p \le \infty$ as well as (probability) measures, including "Dirac delta functions." It will be convenient to write a linear functional $F \in \mathcal{D}'(X)$ symbolically as

$$F(\psi) = \int_X f(x)\psi(x)dx \tag{38}$$

even though there may not exist a pointwise function $f(x)$ that defines the distribution F (e.g. the Dirac distribution). (The f is understood in the context of a limit of a sequence of distributions $F_n(\psi) \in \mathbf{R}$ that, in turn, corresponds to a sequence of test functions $f_n \in \mathcal{D}(X)$. This sequence is guaranteed to exist – see [28] and references therein.)

The goal is to construct an IFS-type fractal transform operator $T : \mathcal{D}'(X) \to \mathcal{D}'(X)$ which, under suitable conditions, will be contractive with respect to a given metric on $\mathcal{D}'(X)$. We begin with the following definition [28]:

DEFINITION 1. *A function $g : \mathbf{R} \to \mathbf{R}$ will be said to satisfy a weak Lipschitz condition on $\mathcal{D}'(X)$ if there exists a $K \ge 0$ such that for all $\psi \in \mathcal{D}(X)$,*

$$\left|\int_X [(g \circ f_1)(x) - (g \circ f_2)(x)]\psi(x)dx\right| \le K\left|\int_X [f_1(x) - f_2(x)]\psi(x)dx\right| \tag{39}$$
$$\forall f_1, f_2 \in \mathcal{D}'(X).$$

Note: If g is affine on $\mathbf{R}$, then it satisfies a weak Lipschitz condition on $\mathcal{D}'(X)$.

Now let $\mathbf{w} = \{w_1, \ldots, w_N\}$ be a set of affine contraction maps in X with contraction factors c_i. Let $\Phi = \{\phi_1, \ldots, \phi_N\}$, be a set of grey level maps satisfying the weak Lipschitz condition on $\mathcal{D}'(X)$ with Lipschitz constants K_i. Associated with the IFS on Distributions (IFSD) $(\mathbf{w}, \Phi)$ is a fractal transform operator $T : \mathcal{D}'(X) \to \mathcal{D}'(X)$ defined as follows. For any $f \in \mathcal{D}'(X)$, the distribution $g = Tf$ will be defined formally by the linear functional

$$\begin{aligned} G(\psi) &= \int_X g(x)\psi(x)dx = \int_X (Tf)(x)\psi(x)dx \\ &= \sum_{i=1} \int_X (\phi_i \circ f \circ w_i^{-1})(x)\psi(x)dx. \end{aligned} \tag{40}$$

We now define the following metric on $\mathcal{D}'(X)$:

$$d_{D'}(f,g) = \sup_{\psi \in \mathcal{D}_1(X)} \left| \int_X f(x)\psi(x)dx - \int_X g(x)\psi(x)dx \right|, \quad \forall f, g \in \mathcal{D}'(X), \tag{41}$$

where $\mathcal{D}_1(X) = \{\psi \in C^\infty(X) \mid \| \psi \|_\infty \leq 1\}$. The metric space $(\mathcal{D}'(X), d_{D'})$ is complete [28]. Then for any $f, g \in \mathcal{D}'(X)$,

$$d_{D'}(Tf, Tg) \leq C_D \, d_{D'}(f,g), \qquad C_D = \sum_{i=1} c_i K_i. \tag{42}$$

If $C_D < 1$, there exists a unique distribution $\bar{u} \in \mathcal{D}'(X)$ such that $T\bar{u} = \bar{u}$. Furthermore, $d_{D'}(T^n u, \bar{u}) \to 0$ as $n \to \infty$ for all $u \in \mathcal{D}'(X)$.

Since the space $\mathcal{D}'(X)$ contains elements of $\mathcal{L}^p(X)$ as well as $\mathcal{M}(X)$, we have now accomplished the sought-after union of IFS methods over function and measure spaces. Note that in the case of IFSP, the K_i are the probabilities p_i. With the condition $\sum_i K_i = 1$, we obtain the usual IFSP contraction factor $C_D = c = \max_i\{c_i\}$, cf. Eq. (9).

3. Generalized fractal transforms. In each of the IFS-type schemes outlined above, an image or target is represented as a point y in an appropriate complete metric space (Y, d_Y). Each element y is some kind of real-valued function/measure/distribution that is supported on the "pixel space" (X, d). Some useful spaces Y have been outlined in the previous section. From a more general perspective, the IFS-type methods over these spaces share a common feature: Given an image or target $u \in Y$, the IFS maps w_i are first used to make N spatially contracted and translated copies $u(w_i^{-1}(x))$ of u. The greyscale values of these copies are then modified by means of appropriate grey level functions ϕ_i (including probabilities p_i for IFSP). These modified copies $\phi_i(u(w_i^{-1}(x)))$ of u are then combined to produce a new function v. The entire process can be summarized as $v = Tu$, where T is a *fractal transform operator* $T : Y \to Y$. Conditions imposed on the w_i and ϕ_i guarantee that T maps Y into itself.

It is also desirable that, subject to additional conditions, T be contractive on (Y, d_Y).

In most, if not all, fractal-based approximation methods, it is assumed that the sets $X_i = w_i(X)$ are *nonoverlapping* (or overlap on a set of zero Lebesgue measure): For almost all $x \in X$, $w_i^{-1}(x)$ exists for only one value $i \in \{1, 2, \ldots N\}$. However, we now wish to consider the more general case of *overlapping*, where a point $x \in X$ can possibly have more than one preimage.

These questions – the common feature of IFS methods to produce modified copies of a function and the question of how to combine these copies in regions of overlap – led to a formulation of *generalized fractal transforms* which is summarized below. (Details are in [29].) The ingredients are as follows:

1. **The base space:** denoted, as above, by (X, d). This is the support of the image functions, hence the space representing the pixels – a compact and connected subset of $\mathbf{R}^n$, typically, $X = [0, 1]^n$ with Euclidean metric. (It may, however, be useful to use $X = \mathbf{R}^n$.)
2. **The IFS component:** For an $N \geq 1$, let $\mathbf{w}$ denote a set of contractive maps, $w_i : X \to X$, $1 \leq i \leq N$. (Or, in the case of local IFS, $w_i : D_i \to R_i$, contractive, where $D_i \subseteq X$. It may also be possible to relax the condition that the w_i be contractive.)
3. **The image function space:** $\mathcal{F}(X) = \{u : X \to R_g\}$, the functions which will represent our images. The **grey level range** $R_g \subset \mathbf{R}$ will denote the range of a particular class of image functions used in a given fractal transform method. (In practical applications, R_g is a bounded subset of $\mathbf{R}^+$.) These image functions will comprise the space (Y, d_Y).
4. **The grey level component:** Associated with the IFS maps $\mathbf{w}$ will be a vector of N functions $\Phi = \{\phi_1, \phi_2, \ldots, \phi_N\}$, $\phi_i : R_g \to R_g$. We may also consider $\phi_i : R_g \times X \to R_g$, i.e., "place-dependent" grey level maps, as discussed earlier.
5. **The fractal components of u** will be given by $f_i \in \mathcal{F}(X)$, $1 \leq i \leq N$, where

$$f_i(x) = \begin{cases} \phi_i(u(w_i^{-1}(x))), & x \in w_i(X), \\ 0, & x \notin w_i(X). \end{cases} \tag{43}$$

6. **The generalized fractal transform of $u \in \mathcal{F}(X)$:** $F_k : [R_g]^k \to R_g$, $1 \leq k \leq N$, where

$$F_k(\mathbf{t}) = F_k(t_1, t_2, \ldots, t_k), \quad t_i \in R_g, \quad 1 \leq i \leq k. \tag{44}$$

The transform F_N defines an operator $T : \mathcal{F}(X) \to \mathcal{F}(X)$ that associates to each image function $u \in \mathcal{F}(X)$ the image function $v = Tu$.

The F_k operators combine the k distinct fractal components $t_i = f_i(x)$ subject to a prescribed set of conditions [11, 28]. (Note that k

depends upon x.) These conditions include (1) permutation symmetry and (2) recursivity which imply that F_2 is an associative binary operation S on $\mathbf{R}^+ \times \mathbf{R}^+$. The choice of the binary operation S depends on the image function space used. The grey level ranges and fractal transforms for IFZS and IFSM are given below:
IFZS: $R_g = [0,1]$. $F_N(x) = \sup_{1\le i\le N}\{f_1(x), f_2(x), \ldots, f_N(x)\}$.
IFSM: $R_g = \mathbf{R}$. $F_N(x) = \sum_{i=1}^N f_i(x)$.
Under suitable conditions that depend upon the space $\mathcal{F}(X)$ and the operator F_k, the generalized fractal transform T is contractive on the space $\mathcal{F}(X)$, implying the existence of a fixed point $\bar{f} \in \mathcal{F}(X)$. The relation $\bar{f} = T\bar{f}$ implies that $\bar{f}$ satisfies the functional equation

$$\bar{f}(x) = F_k(\phi \circ \bar{f} \circ w_1^{-1}(x), \ldots, \phi \circ \bar{f} \circ w_k^{-1}(x)). \tag{45}$$

The presence of $\bar{f}$ on both sides indicates that it possesses a kind of self-similarity, i.e., that it can be expressed as a suitable combination of scaled and modified copies of itself. Eqs. (10) and (15) were already encountered as examples demonstrating this property. Indeed, these equations are special cases of the Bajraktarevic functional equation in Eq. (24).

Remarks on the overlapping property.

1. As mentioned earlier, most, if not all, fractal-based compression methods assume that the fractal components do not overlap. Nonoverlapping simplifies the procedure of "collaging." However, techniques that either simulate or introduce overlapping have been used to reduce the problem of blockiness in fractally encoded images. This was one of the original motivations for fractal-wavelet transforms – to allow overlapping wavelet basis functions (as opposed to the nonoverlapping Haar basis) to perform some kind of smoothing between blocks.
2. The overlapping property was used in formal solutions to the various inverse problems of approximation using IFS-type methods [29]. An infinite set of contraction maps $\mathbf{w} = \{w_1, w_2, \ldots\}$ is employed as a kind of basis set of IFS maps with arbitrarily high degrees of refinement. The sets $X_i = w_i(X)$ necessarily overlap.

3.1. IFS-type operators in other spaces. It may not be obvious how to formulate IFS-type operators in spaces other than those considered above. (It also may not be obvious *why* one would even consider such problems.) We begin with cases that are relatively straightforward. In this section, we assume that both IFS and grey level maps are affine, cf. Eq. (19).

3.1.1. Fractal transforms on faithful representations of functions. Suppose that the elements $y \in Y$ have a faithful representation in another complete metric space (Z, d_Z):

$$\begin{array}{ccc} Y & \xrightarrow{T} & Y \\ \phi \updownarrow & & \updownarrow \phi \\ Z & \xrightarrow{M} & Z \end{array}$$

where ϕ is 1-1 and invertible (and preferably linear, i.e. an isomorphism). Then a fractal transform operator $T : Y \to Y$ induces an equivalent operator $M : Z \to Z$ that is considered to define a fractal transform on Z. Such methods have been useful for the treatment of inverse problems in more suitable spaces. Two noteworthy examples are *moment methods for measures* and *fractal-wavelet transforms for image functions.*

Moment methods for measures: Let (Y, d_Y) be the space $(\mathcal{M}(X), d_M)$ of probability measures on $X = [0,1]$ (for simplicity) with the Monge-Kantorovich metric of Eq. (7). Consider the following space $D(X)$ of moment vectors

$$(46)\quad D(X) = \{\mathbf{g} = (g_0, g_1, g_2, \ldots) \mid g_n = \int_X x^n dx, n = 0, 1, 2, \ldots, \forall \mu \in \mathcal{M}(X)\}.$$

(Note that $g_0 = 1$.) Now define the following metric on $D(X)$: For $\mathbf{u}, \mathbf{v} \in D(X)$,

$$(47)\qquad \bar{d}_2(\mathbf{u}, \mathbf{v}) = \sum_{k=1}^{\infty} \frac{1}{k^2}(u_k - v_k)^2.$$

Then $(D(X), \bar{d}_2)$ is a complete metric space [26]. Now let $(\mathbf{w}, \mathbf{p})$ be an N-map IFSP with associated Markov operator $M : \mathcal{M} \to \mathcal{M}(X)$. There corresponds to M an associated linear operator $A : D(X) \to D(X)$. Moreover, A is contractive in $(D(X), \bar{d}_2)$. As a result, there exists a unique element $\bar{\mathbf{g}} \in D(X)$ such that $A\bar{\mathbf{g}} = \bar{\mathbf{g}}$. (The elements of $\bar{\mathbf{g}}$ are the moments of the invariant measure $\bar{\mu}$ of the IFSP.) Details are given in [26]. A similar procedure may be employed for Fourier transforms of measures [29].

Transforms of $\mathcal{L}^2$ functions. For simplicity of notation, we consider the one-dimensional case, $X = [0,1]$. Let $\{q_n\}_{n=0}^{\infty}$, with $q_0(x) = 1$, denote a complete set of orthonormal basis functions on X. Then for a given $u \in \mathcal{L}^2(X)$,

$$(48)\qquad u(x) = \sum_{k=0}^{\infty} c_k q_k(x),$$

where

$$(49)\qquad c_k = < u, q_k > = \int_X u(x) q_k(x) dx.$$

The infinite vector $\mathbf{c} \in l^2(\mathbf{N})$ is a faithful representation of $u \in \mathcal{L}^2(X)$. Now let $(\mathbf{w}, \Phi)$ denote an N-map affine IFSM on X with with associated operator T and let $v = Tu$. Then

$$v(x) = \sum_{k=0}^{\infty} d_k q_k(x), \tag{50}$$

where

$$d_k = \sum_{l=1}^{\infty} a_{kl} c_l \; + \; e_k, \tag{51}$$

The coefficients

$$a_{kl} = \sum_{i=1}^{N} \alpha_i < q_k, q_l \circ w_i^{-1} >, \qquad e_k = \sum_{i=1}^{N} \beta_i < q_k, I_{w_i(X)} > . \tag{52}$$

define the affine mapping $M : \mathbf{c} \to \mathbf{d}$ associated with the IFSM operator T. If T is contractive in $(\mathcal{L}^p(X), d_p)$ then the mapping $M : l^2(\mathbf{N}) \to l^2(\mathbf{N})$ is contractive in the corresponding metric. (The fixed point $\bar{\mathbf{c}} = M\bar{\mathbf{c}}$ is the vector of Fourier coefficients of the IFSM fixed point function $\bar{u} = T\bar{u}$ in the q_i basis.)

In general, the matrix representation of M is rather full, which is the case for the discrete cosine transform, for example. However, in the special case that the q_k are localized in space, e.g. *wavelets*, many of these matrix elements vanish. The relations between the d_k and the c_l simplify even further when the IFS maps w_i coincide with the dilatations involved in the construction of wavelet bases. This was nicely illustrated by G. Davis for the case of Haar wavelets [16] and then shown more generally in [28]. We discuss the resulting fractal-wavelet transform in the next section.

3.2. Fractal-wavelet transforms. It is instructive to look at the transformation induced on a wavelet expansion tree by a simple affine IFSM. Let $\phi(x)$ denote a scaling function that yields a standard multiresolution approximation of $\mathcal{L}^2(\mathbf{R})$ so that $\psi(x)$ is the corresponding orthogonal mother wavelet function. (For details, see [15, 47].) Then the functions

$$\psi_{ij}(x) = 2^{i/2}\psi(2^i x - j), \quad -\infty < i, j < \infty, \tag{53}$$

form a complete basis of $\mathcal{L}^2(\mathbf{R})$. In the special case of the Haar basis,

$$\phi(x) = I_{[0,1]}(x), \quad \psi(x) = I_{[0,1/2]}(x) - I_{[1/2,1]}(x), \tag{54}$$

We consider functions $f \in \mathcal{L}^2(\mathbf{R})$ admitting the following wavelet expansions:

$$\begin{aligned} f(x) &= b_{00}\phi(x) + \sum_{i=0}^{\infty}\sum_{j=0}^{2i-1} c_{ij}\psi_{ij}(x) \\ &= b_{00}\phi(x) + f_0(x). \end{aligned} \tag{55}$$

The wavelet expansion coefficients are usually displayed in the following manner:

b_{00}							
c_{00}							
c_{10}				c_{11}			
c_{20}		c_{21}		c_{22}		c_{23}	
C_{30}	C_{31}	C_{32}	C_{33}	C_{34}	C_{35}	C_{36}	C_{37}

Each entry or *block* C_{ij} represents a wavelet coefficient tree of infinite length rooted at c_{ij}. The block C_{00} represents the complete wavelet expansion tree of $f_0(x)$. In the case of the Haar basis, f has support $[0,1]$.

Now consider the following two-map affine IFSM:

$$\begin{aligned} w_1(x) &= \frac{1}{2}x, & \phi_1(t) &= \alpha_1 t + \beta_1, \\ w_2(x) &= \frac{1}{2}x + \frac{1}{2}, & \phi_2(t) &= \alpha_2 t + \beta_2, \end{aligned} \tag{56}$$

and let $g = Tf$. Then

$$\begin{aligned} g(x) = \alpha_1 b_{00}\phi(2x) + \sum_{i,j} c_{ij}\psi_{ij}(2x) + \beta_1 I_{[0,1/2]}(x) \\ + \alpha_2 b_{00}\phi(2x-1) + \sum_{i,j} c_{ij}\psi_{ij}(2x-1) + \beta_2 I_{[1/2,1]}(x). \end{aligned}$$

It is convenient to write $g(x)$ in the form

$$g(x) = \theta(x) + g_0(x), \tag{57}$$

where

$$\theta(x) = [\alpha_1 b_{00} + \beta_1]I_{[0,1/2]}(x) + [\alpha_2 b_{00} + \beta_2]I_{[1/2,1]}(x). \tag{58}$$

In the Haar basis, the components of $\theta(x)$ lie only in the top two entries of the wavelet tree occupied by the scaling and mother wavelet components. The function $g_0(x)$ has the wavelet expansion

(59)

0	
0	
$\frac{1}{\sqrt{2}}\alpha_1 C_{00}$	$\frac{1}{\sqrt{2}}\alpha_2 C_{00}$

.

The IFS operator T is seen to induce an IFS-type operation on the wavelet coefficient trees: The two blocks C_{10} and C_{11} of the wavelet expansion of f are replaced by scaled copies of C_{00}. In addition, the the b_{00} and c_{00} entries representing $\theta(x)$ are modified appropriately.

In the Haar basis, such a scaling and copying of higher subtrees onto lower subtrees is also produced by appropriate "local IFSM" that map

dyadic domain blocks to small range blocks. This connection, made possible by the nonoverlapping nature of Haar wavelet basis functions, has been discussed in in [16, 29]. From this connection, a number of researchers independently defined *discrete fractal-wavelet transforms* that performed scaling and copying operations on wavelet coefficient trees for generalized (non-Haar) systems of compactly supported wavelets ([16, 29, 40, 59, 62] to name a few). In these cases, the supports of contiguous wavelet functions overlap. The motivation to devise such transforms was an attempt to reduce the blockiness exhibited by usual fractal (= Haar) block-encoding schemes.

Indeed, such fractal-wavelet (FW) transforms represent an interesting IFS-type of operation on wavelet subtrees. Working backwards, however, the connection between these operations and the transformations they induce in function space are not quite as straightforward as in the Haar case. In the case of nonperiodized wavelets, the discrete FW transform is equivalent to a recurrent IFSM [51]. In the case of *periodized* wavelets, which have been used in image processing, the connection is an even more complicated form of IFSM. (The connection in the case of *biorthogonal wavelets* that are employed in most current image processing applications has not yet been considered.)

We illustrate the basic idea of FW transforms with a simple example taken from [51]. Consider the following FW transform with four block maps:

$$W_1 : C_{10} \to C_{20},\ W_2 : C_{11} \to C_{21},\ W_3 : C_{10} \to C_{22},\ W_4 : C_{11} \to C_{23}, \tag{60}$$

with associated multipliers α_i, $1 \le i \le 4$. Diagramatically,

$$M\ :\ C_{00} \to \begin{array}{|c|c|c|c|} \hline \multicolumn{4}{|c|}{c_{00}} \\ \hline \multicolumn{2}{|c|}{c_{10}} & \multicolumn{2}{c|}{c_{11}} \\ \hline \alpha_1 C_{10} & \alpha_2 C_{11} & \alpha_3 C_{10} & \alpha_4 C_{11} \\ \hline \end{array} \quad . \tag{61}$$

Note that the coefficient b_{00} remains unchanged in this FW operation – for this reason its contribution is ignored. Now iterate this process, assuming that it converges to a limit $\bar{C}_{00}$ which represents the wavelet expansion of a function $\bar{u} \in \mathcal{L}^2$. (A look at the sequence of wavelet coefficients generated by the iteration shows that the condition $|\alpha_i| < \frac{1}{\sqrt{2}}$ is sufficient to guarantee the existence of such a limit.) Then

$$\bar{u} = c_{00}\psi_{00} + \bar{v}. \tag{62}$$

The function $\bar{v}$ admits the wavelet expansion

$$\begin{array}{|c|c|c|c|} \hline \multicolumn{2}{|c|}{c_{10}} & \multicolumn{2}{c|}{c_{11}} \\ \hline \alpha_1 C_{10} & \alpha_2 C_{11} & \alpha_3 C_{10} & \alpha_4 C_{11} \\ \hline \end{array} \quad . \tag{63}$$

Since $< \psi_{10}, \psi_{11} >= 0$, etc., we may write

$$\bar{v} = \bar{v}_1 + \bar{v}_2 , \tag{64}$$

where the components $\bar{v}_i$ satisfy the relations

$$\begin{aligned} \bar{v}_1(x) &= c_{10}\psi_{10}(x) + \alpha_1\sqrt{2}\bar{v}_1(2x) + \alpha_2\sqrt{2}\bar{v}_2(2x) \\ \bar{v}_2(x) &= c_{11}\psi_{11}(x) + \alpha_3\sqrt{2}\bar{v}_1(2x-1) + \alpha_4\sqrt{2}\bar{v}_2(2x-1). \end{aligned} \tag{65}$$

These equations define a *vector IFSM with condensation.* The vector $\bar{v}$ is composed of the orthogonal components $\bar{v}_1$ and $\bar{v}_2$ that satisfy the above fixed point relations.

Note that the contractive IFS maps w_{ij} are mappings from the *entire* base space X into itself and not local IFS maps. What appeared to be a local transform in wavelet coefficient space is a vector IFSM in the base space. (Again, in the special case of the nonoverlapping Haar wavelets, the above IFSM may be written as a local IFSM.) The locality of the block transform has been passed on to the orthogonal components $\bar{v}_1$ and $\bar{v}_2$ of the function $\bar{v}$. These components may be considered as "nonoverlapping" elements of a vector. We refer the reader to [51] for a more detailed discussion of such fractal-wavelet transforms and the operations that they induce in function space.

3.2.1. Generalized (2D) fractal-wavelet transforms for images. Two-dimensional fractal-wavelet transforms involve mappings of "parent" quadtrees of wavelet expansions to lower "child" quadtrees as was done for binary trees in the one-dimensional case. In the discussion that follows, we assume the standard construction of orthonormal wavelet bases in $\mathcal{L}^2(\mathbf{R}^2)$ using suitable tensor products of 1D basis functions. Once again we assume the existence of a scaling function $\phi(x)$ and its corresponding (orthogonal) mother wavelet function $\psi(x)$ that give rise to a 1D multiresolution analysis [15, 47]. Let

$$\phi_{ij}(x) = 2^{i/2}\phi(2^i x - k), \quad \psi_{ij}(x) = 2^{i/2}\psi(2^i x - k), \tag{66}$$

and define the following orthogonal subspaces spanned by appropriate tensor products of these functions:

$$\begin{aligned} V_k^0 &= \text{span}\{\phi_{kij}(x,y) = \phi_{ki}(x)\phi_{kj}(y),\ 0 \le i,j \le 2^k - 1\} \\ W_k^h &= \text{span}\{\psi_{kij}^h(x,y) = \phi_{ki}(x)\psi_{kj}(y),\ 0 \le i,j \le 2^k - 1\} \\ W_k^v &= \text{span}\{\psi_{kij}^v(x,y) = \psi_{ki}(x)\phi_{kj}(y),\ 0 \le i,j \le 2^k - 1\} \\ W_k^d &= \text{span}\{\psi_{kij}^d(x,y) = \psi_{ki}(x)\psi_{kj}(y),\ 0 \le i,j \le 2^k - 1\}. \end{aligned} \tag{67}$$

The superscripts h, v and d stand for *horizontal*, *vertical* and *diagonal*, respectively [15, 47].

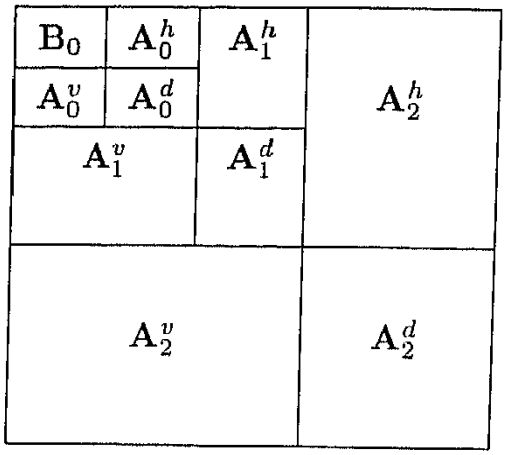

FIG. 3. *Matrix arrangement of two-dimensional wavelet coefficient blocks.*

We consider functions $u(x,y)$ admitting the following wavelet expansions:

$$u(x,y) = b_{000}\phi_{000}(x,y) + \sum_{k=0}^{\infty}\sum_{i=0}^{2^k-1}\sum_{j=0}^{2^k-1}[a^h_{kij}\psi^h_{kij}(x,y)+a^v_{kij}\psi^v_{kij}(x,y)+a^d_{kij}\psi^d_{kij}(x,y)]. \tag{68}$$

The wavelet expansion coefficients a^λ_{kij} are conveniently arranged in a standard fashion ([15, 47]) as shown Figure 3. Each of the blocks A^h_k, A^v_k, A^d_k, $k \geq 0$, contains 2^{2k} coefficients $a^h_{kij}, a^v_{kij}, a^d_{kij}$, respectively. The three collections of blocks

$$A^h = \bigcup_k^\infty A^h_k, \quad A^v = \bigcup_k^\infty A^v_k, \quad A^d = \bigcup_k^\infty A^d_k, \tag{69}$$

comprise the fundamental *horizontal*, *vertical* and *diagonal* quadtrees of the coefficient tree.

Now consider any wavelet coefficient a^λ_{kij}, $\lambda \in \{h,v,d\}$ in this matrix and the unique (infinite) quadtree with this element as its root. We shall denote this quadtree as A^λ_{kij}. In the Haar case, for a fixed set of indices $\{k,i,j\}$ the three quadtrees A^h_{kij}, A^v_{kij} and A^d_{kij} correspond to the *same spatial block* of the function or image.

Two-dimensional fractal-wavelet transforms involve mappings of "parent" quadtrees of wavelet expansions to lower "child" quadtrees. For simplicity in presentation and notation, we consider a particular case in which the roots of all parent quadtrees appear in a given block and the roots of all child quadtrees appear in another given block. The method is easily modified for other schemes, for example, quadtree partitioning.

Select two integers, the parent and child levels, k_1^* and k_2^*, respectively, with $0 \leq k_1^* < k_2^*$. For each possible index, $0 \leq i,j \leq 2^{k_2^*} - 1$, define the three sets of affine block transforms:

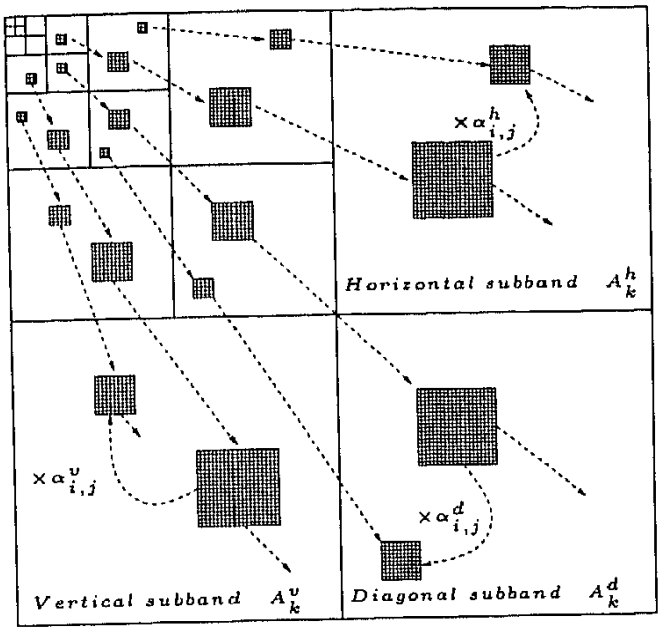

FIG. 4. *The 2-dimensional fractal-wavelet transform.*

$$W_{ij}^{\lambda} : A^{\lambda}_{k_1^*, i^{\lambda}(i,j), j^{\lambda}(i,j)} \to A^{\lambda}_{k_2^*, i, j},$$
$$A^{\lambda}_{k_2^*,i,j} = \alpha^{\lambda}_{ij} A^{\lambda}_{k_1^*, i^{\lambda}(i,j), j^{\lambda}(i,j)}, \quad \lambda \in \{h, v, d\}. \tag{70}$$

Notice how the child quadtrees at level k_2^* are replaced by scaled copies of parent quadtrees from level k_1^*. The procedure is illustrated in Figure 4. These block transforms comprise a unique FW operator M. The use of the indices i^h, j^h, etc. emphasizes that the parent quadtrees corresponding to a given set of child quadtrees $A^h_{k_2^*,i,j}$, $A^v_{k_2^*,i,j}$ and $A^d_{k_2^*,i,j}$ need *not* be the same. As well, the scaling coefficients $\alpha^h_{ij}, \alpha^v_{ij}$ and α^d_{ij} can be *independent*. The "fractal code" associated with the operator M consists of the following:

1. The parent-child index pair (k_1^*, k_2^*).
2. The scaling coefficient b_{000} in block $\mathbf{B_0}$ along with the wavelet coefficients in blocks A^{λ}_k, $\lambda \in \{h, v, d\}$ for $1 \le k \le k_2^* - 1$. $4^{k_2^*}$ coefficients.
3. The scaling factors α^{λ}_{ij} and parent block indices, $(i^{\lambda}(i,j),\ j^{\lambda}(i,j))$, for all elements a^{λ}_{ij} in each of the three blocks $A^{\lambda}_{k_2^*}$. Total number of parameters: (i) $3 \cdot 4^{k_2^*}$ scaling factors, (ii) $2 \cdot 3 \cdot 4^{k_2^*}$ indices.

It has been shown [64] that, under certain conditions, the fractal-wavelet transform M is contractive in an appropriate complete metric space (l_2 square summable sequences) of wavelet coefficients. For the special transform given in Eq. (70), contractivity is guaranteed when

$$c_Q = 2^{k_2^* - k_1^*} \max_{\lambda,i,j} |\alpha^{\lambda}_{ij}| \ < \ 1, \tag{71}$$

where $\lambda \in \{h, v, d\}$ and $0 \le i, j \le 2^{k_2^*} - 1$. The condition $c_Q < 1$ guarantees the existence of a unique fixed point wavelet coefficient tree $\bar{\mathbf{A}} = M\bar{\mathbf{A}}$.

From the definition of M, $\bar{\mathbf{A}}$ is a union of scaled copies of its subtrees, a kind of local self-similarity property.

The wavelet tree $\bar{\mathbf{A}}$ may be generated by iteration of M. Some simple examples are presented in [64]. In practical applications, e.g. images, one may begin with a wavelet coefficient matrix – call it $\mathbf{C}_0$ – containing the "fractal code" scaling and wavelet coefficients in 2) above, with all other blocks A_k^λ, $k \geq k_2^*$ being zeros. In the iteration procedure $\mathbf{C}_{n+1} = M\mathbf{C}_n$, each application of M produces an additional level of blocks, representing an additional degree of refinement of the function u in terms of its wavelet expansion. Note that such an iteration of M essentially produces a geometric-type *extrapolation* of the wavelet coefficients of the base matrix $\mathbf{C}_0$, involving products of the scaling coefficients a_{kij}^λ. We shall return to this idea in a later section.

In standard FW schemes [16, 40, 59, 62], common parents and common scaling factors are used for the various subbands, that is:

$$\begin{aligned} i^h(i,j) &= i^v(i,j) = i^d(i,j) \\ j^h(i,j) &= j^v(i,j) = j^d(i,j) \\ \alpha_{ij}^h &= \alpha_{ij}^v = \alpha_{ij}^d. \end{aligned} \tag{72}$$

In other words, the h, v and d subbands are *not* treated independently.

The FW transform M induces an equivalent mapping in function space. As in the case of 1D FW transforms, such mappings are generally a kind of vector IFSM operator with condensation functions, once again performing scaling and mixing operations among orthogonal components of a function. Some examples are given in [64].

3.3. IFS on integral operators. In [25], the "parallel" space (Z, d_Z) was considered to represent an appropriate space of function transforms, namely, *integral transforms*. There are at least two motivations to consider integral transforms:

1. In many cases, e.g. MRI, blurred images, the data that we seek to represent or compress is the result of an integral transform on some function space.
2. It may be more convenient to work in certain spaces of integral transforms. For example, Lebesgue transforms of normalized non-negative $\mathcal{L}^1$ functions are nondecreasing and continuous functions. They may be easier to work with, especially in the sense of approximability.

In this section we let $\mathcal{S} : \mathcal{F} \to \mathcal{G}$ denote an integral transform with kernel $K : X \times \mathbf{R} \to \mathbf{R}$,

$$\widehat{f}(s) = (\mathcal{S}f)(s) = \int_X K(t,s)f(t)\, dt. \tag{73}$$

We shall also write this transform in inner product form as $\mathcal{S}f =< K, f >$.

Let T be an affine IFSM operator as defined in Eq. (19). For an $f \in \mathcal{L}^1(X)$, let $g = Tf$. Then the transform $\widehat{g} = \mathcal{S}(g)$ is given by

$$\begin{aligned}\widehat{g}(s) &= \int_X K(t,s) \sum_{i=1}^{N} \left[\alpha_i f(w_i^{-1}(t)) + \beta_i\right] I_{X_i}(t)\, dt \\ &= \sum_{i=1}^{N} \alpha_i \int_{X_i} K(t,s) f(w_i^{-1}(t))\, dt \;+\; \sum_{i=1}^{N} \beta_i \int_{X_i} K(s,t) dt \\ &= \sum_{i=1}^{N} \alpha_i c_i \int_X K(c_i u + a_i, s) f(u)\, du \;+\; \widehat{\beta}(s),\end{aligned} \tag{74}$$

where

$$\widehat{\beta}(s) = \sum_{i=1}^{N} \beta_i \widehat{I_{X_i}}(s). \tag{75}$$

(Note that $\widehat{\beta}(s)$ depends only on the β_i – and, of course, the X_i – but not on f.)

Eq. (74) may be written in the form

$$< K, Tf > = < T^\dagger K, f > \;+\; L(s), \qquad f \in \mathcal{F}. \tag{76}$$

The operator $T^\dagger$ may be interpreted as a kind of "adjoint" fractal operator on the kernel K,

$$(T^\dagger K)(t,s) = \sum_{i=1}^{N} \alpha_i c_i K(c_i t + a_i, s), \tag{77}$$

and L as a kind of condensation function. However, the dilations in the spatial variable produced by $T^\dagger$ in the above equation represent *expansions*. In contrast to IFSM fractal transforms on functions, the transform K is tiled with expanded copies of itself. (This was well known for the case of Fourier/Laplace transforms [29, 33].)

In an effort to express the integrals in Eq. (74) involving K as *bona fide* integral transforms of f, one may postulate that K must satisfy a general functional relation of the form

$$K(c_i u + a_i, s) = C(c_i, a_i, s) K(u, \zeta(c_i, a_i, s)), \quad \forall u \in X, \quad i = 1, 2, \ldots, N. \tag{78}$$

This equation may be considered in a number of ways, including:

1. A functional relation between the kernel K, the constant C and scaling function ζ,
2. A functional equation in the unknown functions K and ζ, given C,
3. A functional equation in the unknown functions C and ζ, given K.

As in the case of differential equations, the solution of functional equations requires initial conditions. In addition, however, an admissible space of functions in which solutions are sought must also be specified. Some simple results are presented in [25]. In addition, the familiar cases of integral transforms – Fourier, wavelet, Lebesgue, as well as moments of measures – are shown to be covered by the above general formalism.

3.4. IFS on vector-valued measures. In [52], IFS-type transforms over self-similar vector-valued Borel measures were defined. This method, motivated by [33], permits the construction of tangent and normal vector measures to planar fractal curves. In this way, line integrals of smooth vector fields over planar fractal curves may be defined. This leads to a formulation of Green's theorem and the Divergence theorem for planar regions bounded by fractal curves.

Very briefly, let w_i be contractive IFS maps on X and $p_i > 0$ be associated real numbers without the restriction that $\sum_i p_i = 1$. Also let R_i denote linear operators on $\mathbf{R}^n$. We let $\mathcal{M}^n(X, \mathbf{R}^n)$ denote the set of vector-valued measures on the Borel sigma field of X with values in $\mathbf{R}^n$. Then associated with the "IFSVVM" $(\mathbf{w}, \mathbf{p}, \mathbf{R})$ is an operator $T : \mathcal{M}^n(X, \mathbf{R}^n) \to \mathcal{M}^n(X, \mathbf{R}^n)$ with action as follows:

$$(T\mu)(B) = \sum_i p_i R_i \mu(w_i^{-1}(B)) \tag{79}$$

for all Borel sets $B \subset X$. This operator is an obvious modification of the scalar IFSP Markov operator M in Eq. (8). F. Mendivil has reported on this work at this conference.

4. Inverse problems for generalized fractal transforms. We consider target functions or images to be elements of an appropriate complete metric space (Y, d_Y). The underlying idea in fractal compression is the approximation, to some suitable accuracy, of a target $y \in Y$ by the fixed point $\bar{y}$ of a contraction mapping $f : Y \to Y$. It is then f which is stored in computer memory. By Banach's Fixed point theorem the unique fixed point $\bar{y}$ may be generated by iteration of f, using an arbitrary "seed" image $y_0 \in Y$.

Naturally, most effort in this area has focussed on the compression of digital images: For a given accuracy (typically $\mathcal{L}^2$ error) find a fractal transform f such that the parameters defining it require the least amount of computer memory after quantization and entropy coding. As in any compression scheme, there is a competition between accuracy of approximation and the compression/reduction of data. The various theoretical and practical aspects of fractal image compression have been covered very well in the books by Fisher [23], Barnsley and Hurd [9] and Lu [46].

In practical fractal image coding, the parameter space $\mathcal{P}$ of feasible, quantized fractal codes corresponding to a given scheme (partitioning, choice of domain pools, etc.) is discrete and finite. As such, there

exists a minimum value of the *attractor error* $d_Y(y, \bar{y}(p))$ for some $p \in \mathcal{P}$. Since digital images are described to only a finite resolution and since, for practical purposes, fractal coding schemes will not employ partitions comprised solely of single pixels, the attractor error will generally be nonzero. However, it is a tedious, if not typically intractable, procedure to determine such optimal codes. This is the reason, as is well known, that fractal coding schemes generally employ "collage coding". In fact, Ruhl and Hartenstein [57] have shown that optimal fractal coding is an NP-hard problem.

In what follows, we outline the work to establish a more general theory of fractal-based approximation in continuous spatial and grey level variables, as done in [26–29]. The central theme is the mathematical solution to the following *formal inverse problem of approximation by fixed points of contractive operators*:

Define the set of contraction maps on (Y, d_Y) as follows:

$$Con(Y) = \{f : Y \to Y \mid d_Y(f(x), f(y)) \leq c_f d_Y(x, y) \quad \forall x, y \in Y,\ c_f \in [0, 1)\}. \tag{80}$$

Then:

> Given a "target" $y \in Y$ and an $\epsilon > 0$, find a map $f_\epsilon \in Con(Y)$ such that $d_Y(y, \bar{y}_\epsilon) < \epsilon$, where $f_\epsilon(\bar{y}_\epsilon) = \bar{y}_\epsilon$.

In other words, we look for conditions that guarantee that a target $y \in Y$ can be approximated to arbitrary accuracy by the fixed point of a contraction map $f \in Con(Y)$.

There are three important mathematical results which provide the basis for fractal transform methods and fractal-based compression. It is worthwhile to list them here.

1. **Banach fixed point theorem for contraction maps** [3]:

THEOREM 1. *Let (Y, d_Y) be a complete metric space. Suppose there exists a mapping $f \in Con(Y)$ with contractivity factor $c \in [0, 1)$. Then there exists a unique $\bar{y} \in Y$ such that $f(\bar{y}) = \bar{y}$. Moreover, for any $y \in Y$, $d_Y(f^n(y), \bar{y}) \to 0$ as $n \to \infty$.*

2. **Continuity of fixed points with respect to contraction maps** [13]:

THEOREM 2. *Let (Y, d_Y) be a compact metric space and $Con(Y)$ be an appropriate space of contraction maps on Y with the following metric:*

$$d_{Con(Y)}(f, g) = \sup_{y \in Y} d_Y(f(y), g(y)), \quad \forall f, g \in Con(Y). \tag{81}$$

Let $f, g \in Con(Y)$ with fixed points $\bar{y}_f$ and $\bar{y}_g$, respectively. Then

$$d_Y(\bar{y}_f, \bar{y}_g) \leq \frac{1}{1 - \min(c_f, c_g)} d_{Con(Y)}(f, g), \tag{82}$$

where c_f, c_g denote the contractivity factors of f and g, respectively.

This result is a generalization of Barnsley's "continuity with respect to a parameter" result [5]. It was used to derive continuity properties of IFS attractors and IFSP invariant measures [13] as well as IFZS attractors [24]. *Although never stated explicitly, fractal compression algorithms depend on this property since an optimization of the approximation of a target involves the variation of parameters that define fractal transform operators.* (In special cases, IFS attractors are also differentiable with respect to fractal parameters [65].)

3. **"Collage theorem"** [8]:

THEOREM 3. *Let (Y, d_Y) be a complete metric space and let $f \in Con(Y)$ with contractivity factor $c_f \in [0, 1)$. Then for any $y \in Y$,*

$$d_Y(y, \bar{y}) \leq \frac{1}{1 - c_f} d_Y(y, f(y)), \tag{83}$$

where $\bar{y}$ is the fixed point of f.

This result follows from Banach's theorem by using a simple triangle inequality. It appears as a remark to Banach's theorem in [60]. In fact, another manipulation of the triangle inequality involving y, $f(y)$ and $\bar{y}$ yields the following interesting result:

4. **"Anti-Collage theorem"** [65]:

THEOREM 4. *Assume the conditions of Theorem 3. Then for any $y \in Y$,*

$$d_Y(y, \bar{y}) \geq \frac{1}{1 + c_f} d_Y(y, f(y)), \tag{84}$$

where $\bar{y}$ is the fixed point of f.

Given a suitable space $\mathcal{P}$ of acceptable parameters that define contraction mappings $f \in Con(Y)$, it is generally a tedious procedure - even for "non-fractal" problems (cf. Section 5) - to determine the best fixed point approximation to a target, that is, the mapping f_{opt} whose fixed point $\bar{y}_{opt}$ yields the smallest possible *attractor error* $d_Y(y, \bar{y})$. For this reason most, if not all, fractal coding methods rely on a reformulation of the inverse problem made possible by the Collage theorem. Instead of searching for contraction maps f the fixed points $\bar{y}$ of which lie close to a target y (and most probably having to compute $\bar{y}$ by iteration), we look for maps f that map y close to itself. The reformulated inverse problem becomes:

> Given a target $y \in Y$ and a $\delta > 0$, find a map $f_\delta \in Con(Y)$ such that $\Delta = d_Y(y, f_\delta(y)) < \delta$.

The term Δ is often referred to as the *collage distance*. From the Collage theorem, the fixed point $\bar{y}_\delta$ of f_δ will lie within a multiple of δ.

Interestingly, the Collage and Anti-Collage theorems provide upper and lower bounds to the approximation of y by $\bar{y}$ in terms of the collage distance Δ. A nonzero collage distance keeps the error $d_Y(y, \bar{y})$ away from zero (unless, of course, $y = \bar{y}$), a consequence of the triangle formed by y, $f(y)$ and $\bar{y}$.

4.1. Solutions to inverse problems for measures and functions. The basic strategy in solving the inverse problem lies in working with an infinite set $\mathbf{w} = \{w_1, w_2, \ldots,\}$ of fixed (affine) IFS contraction maps that satisfy refinement conditions for the particular metric spaces (Y, d_Y) concerned. (The reader is referred to the appropriate references for details regarding these refinement conditions.) A useful set of IFS maps on [0,1] that satisfy such refinement conditions for both measure and function approximation is

$$w_{ij}(x) = \frac{1}{2^i}(x + j - 1), \quad i = 1, 2, \ldots, \quad j = 1, 2, \ldots, 2^i. \tag{85}$$

(The extension to $[0, 1]^2$ is straightforward.)

The two major inverse problems which have been considered are

1. **Measures:** on the space $(Y, d_Y) = (\mathcal{M}(X), d_H)$ [26],
2. **Functions:** on the space $(Y, d_Y) = (\mathcal{L}^p(X), d_p)$ [27].

We now summarize the main points in the solutions to these problems. In both cases, we choose N-map truncations of $\mathbf{w}$, $\mathbf{w}^N = \{w_1, \ldots, w_N\}$, in order to construct either

1. **Measures:** N-map affine IFSP $(\mathbf{w}^N, \mathbf{p}^N)$, where

$$\mathbf{p}^N \in \Pi^N = \{(p_1^N, \ldots, p_N^N) \mid p_i^N \geq 0, \sum_{i=1}^{N} p_i^N = 1\} \subset \mathbf{R}^{2N}(\text{compact}). \tag{86}$$

$\mathcal{P} = \Pi^N$ is the feasible set of probability vectors for N-map IFSP. Each point in Π^N defines a fractal transform (Markov) operator T^N in $Con(\mathcal{M}(X))$.

2. **Functions:** N-map affine IFSM $(\mathbf{w}^N, \Phi_N)$, where $\Phi^N = \{\phi_1^N, \ldots, \phi_N^N\}$,

$$\phi_i^N(t) = \alpha_i^N t + \beta_i^N, \quad (\alpha^N, \beta^N) \in \Pi^{2N} \subset \mathbf{R}^{2N}(\text{compact}). \tag{87}$$

$\mathcal{P} = \Pi^{2N}$ is the feasible set of grey level map parameters for N-map affine IFSM. Each point in Π^{2N} defines a fractal transform operator $T^N \in Con(\mathcal{L}^p(X))$.

In both approximation problems, the idea is very simple. For functions, we exploit the property that the set of simple functions on X is dense in $\mathcal{L}^2(X)$. For measures, we exploit the property that the set of measures with finite support is dense in $\mathcal{M}(X)$. One proceeds as follows. Given a target $y \in Y$, then for an $N > 0$, find the minimum collage distance

$$\Delta_{\min}^N = \min_{\mathcal{P}(N)} d_Y(y, T^N y). \tag{88}$$

(The minimum exists due to the compactness of the feasible sets Π^N and Π^{2N}.) In both cases, we have the important result:

$$\Delta_{\min}^N \to 0 \text{ as } N \to \infty. \tag{89}$$

This guarantees the existence of solutions to the formal inverse problems for measure and function approximation. The solution of these problems is equivalent to the following result:

COROLLARY 1. *Given a fixed set of affine IFS maps* $\mathbf{w} = \{w_1, w_2, \ldots\}$ *satisfying the appropriate refinement conditions for* (Y, d_Y) *being either (a)* $(\mathcal{M}(X), d_H)$ *(measures) or (b)* $(\mathcal{L}^p(X), d_p)$ *(functions), then the set of all fixed point attractors for (a)* $(\mathbf{w}^N, \mathbf{p}^N)$ *or (b)* $(\mathbf{w}^N, \Phi^N)$, *respectively,* $N = 1, 2, \ldots$, *is dense in* (Y, d_Y).

In both the measure and function approximation problems, the minimization of the squared collage distance $(\Delta^N)^2$ in the appropriate metric becomes a quadratic programming problem in the probability/grey level map vector with constraints. Some numerical results of these algorithms are presented in [26, 27, 29].

4.2. "Direct" methods of fractal image compression. The majority of efforts in fractal image compression have been involved with the image functions: Given a target image function u, find a fractal transform operator T that minimizes the collage distance $\| u - Tu \|$, usually in the $\mathcal{L}^2$ norm. We classify such methods that operate in the image function space (Y, d_Y) as *direct* methods. The formal solutions outlined in the previous section also represent direct methods.

It is not possible here to provide even a summary of the various methods that have been devised to perform fractal image compression. For this, we refer reader to books that have been dedicated to the subject – [22, 23, 9, 46] – as well as the marvellous repository of research papers on fractal image compression stored at the Leipzig Fractal Image Compression website `http://www.informatik.uni-leipzig.de/cgip/`.

4.2.1. A note on the suboptimality of collage coding. As far as practical fractal image coding is concerned, it is well known that collage coding is suboptimal. Collage coding is a *greedy algorithm* that seeks to solve the fractal image approximation problem in one scan of the image. Suppose that for a target $y \in Y$, the contraction mapping T_c minimizes the *collage error* $d_Y(y, Ty)$. Then the fixed point $\bar{y}_c$ of T_c does *not* necessarily minimize the *attractor error* $d_Y(y, \bar{y})$, i.e. $\bar{y}_c$ is not necessarily $\bar{y}_{opt}$ defined earlier. In fact, it has been shown [57] that the ratio of collage error to the optimal attractor error can be arbitrarily large.

In [65] some systematic methods to perform *attractor optimization* – finding better fixed point approximations to a target y than the collage attractor $\bar{y}_c$ – were examined, following earlier work by others (see references in [65]). Often, one begins with the collage attractor, performing a local search in parameter space $\mathcal{P}$ in an attempt to lower the attractor error. In this study, it was shown that affine IFSM attractors are differentiable functions of the grey level map parameters α_i and β_i. This permitted the use of gradient descent methods in the search. Unfortunately, there was no advantage in employing such methods – simple hill-climbing

algorithms (e.g. Nelder-Mead) would yield virtually identical results, often with much less computational expense. (The computation of the gradients is quite complicated. Even for the simple partitioning scheme employed in our study, the partial derivatives comprised a vector IFSM.) As well, the improvements over the collage error were very low, on the order of 0.5 dB in PSNR. A limitation of such methods is that they keep the parent-child assignments fixed.

4.3. Indirect methods in the inverse problem. *Indirect* methods involve the formulation of the inverse problem in (Y, d_Y) as an equivalent inverse problem in the faithful representation (Z, d_Z). Two noteworthy examples that will be summarized below are: (1) Formulating the inverse problem for measure approximation using IFSP as an inverse problem in "moment space," and (2) Fractal-wavelet transforms.

4.3.1. Moment matching methods for measure approximation. Much of the early work on the IFS approximation of measures was based on the idea of matching the moments of IFSP invariant measures as closely as possible to the moments of the target measure. The motivation lay in the fact that moments of invariant measures of affine IFSP can be computed in terms of the IFS parameters in a recursive manner [6]. Indeed, it was in [6] that such an indirect inverse method was first performed. Barnsley and Demko estimated the non-zero (complex) moments g_2 and g_4 of the "twin dragon" region in the plane, assuming a normalized ($g_0 = 1$) and uniform (2D Lebesgue) measure over the region. Using the two complex IFS maps $w_1(z) = sz + (1-s)a$, $w_2(z) = sz - (1-s)a$, they computed the expressions for these two moments in terms of the IFS parameters s and a, with $p_1 = p_2 = \frac{1}{2}$. Matching these moments to the estimated target measure moments yielded approximate values of s and a that were reasonably close to the correct values.

Indeed, this inspired a number of works that applied some form of moment matching along the following lines: Given a target measure $\nu \in \mathcal{M}(X)$ (let $X = [0,1]$ for simplicity) with moments $g_n = \int_X x^n dx, n = 0, 1, 2, \ldots$, find an IFS invariant measure $\bar{\mu}$ whose moments $\bar{g}_n = \int_X x^n dx$ are "close" to the g_n for $n = 1, 2, \ldots, M$, where $M > 1$. (We do not list all relevant references here but rather refer the reader to [63] for a list of important early papers.) Early investigations sought to minimize a sum of squared distances between target and IFS moments. In most cases, methods were devised to find optimal affine IFS maps w_i and associated probabilities p_i.

In [26], the moment matching problem was formulated over the metric space $(D(X), \bar{d}_2)$ of infinite moment vectors introduced in Section 3.1.1. Recall from that section that each N-map IFSP defines a linear operator $A : D(X) \to D(X)$ that is *contractive* in $(D(X), \bar{d}_2)$. Hence the following *Collage theorem for Moments*:

THEOREM 5. *Let (X,d) be a compact metric space and $\mu \in \mathcal{M}(X)$ with moment vector $\mathbf{g} \in D(X)$. Let $(\mathbf{w},\mathbf{p})$ be an N-map IFSP with contractivity factor $c \in [0,1)$. Let $\nu \in \mathcal{M}(X)$ with associated moment vector $\mathbf{h} \in D(X)$. Then*

$$\bar{d}_2(\mathbf{g},\bar{\mathbf{g}}) \le \frac{1}{1-c}\bar{d}_2(\mathbf{g},A\mathbf{g}), \tag{90}$$

where $\bar{\mathbf{g}}$ is the moment vector corresponding to $\bar{\mu}$, the invariant measure of the IFSP $(\mathbf{w},\mathbf{p})$.

In [26], fixed sets of affine IFS maps w_i were used. The inverse problem then reduces to the determination of probabilities p_i that minimize the moment collage distance $\bar{d}_2(\mathbf{g},A\mathbf{g})$. The minimization of this distance in $D(X)$ is a quadratic programming problem with linear constraints on the p_i.

4.3.2. Image compression using fractal-wavelet transforms. Section 3.2 outlined the mathematical apparatus necessary for performing "indirect" methods for image approximation/compression using FW transforms. The space (Z,d_Z) is an appropriate l^2 space of wavelet expansion coefficients and $Con(Z)$ consists of FW transforms M defined by the scaling coefficients α^{λ}_{kij} (as well as the associated parent-child indices). In the collage FW coding of a target image u with wavelet coefficient matrix $\mathbf{A}$, one seeks to find an FW transform M that maps $\mathbf{A}$ as close as possible to itself so that the fixed point $\bar{\mathbf{A}}$ will be a good approximation to $\mathbf{A}$. Of course, in practical applications of image compression, one is concerned with the competition between increased accuracy of the approximation and its "cost" in terms of computer storage of the fractal code.

We have already mentioned that FW transforms were devised independently by a number of workers some time ago (e.g. [16, 40, 59, 62]). Generally, the motivation was to reduce the blockiness that plagued fractally coded images. Since that time, fractal-wavelet coders have been shown [10, 45, 30] to be able to demonstrate rate-distortion performances that can match state-of-the-art (at the time) wavelet coders such as SPIHT [58]. We also mention that FW transforms have been used quite effectively for the compression of one-dimensional audio signals [66].

Our work on fractal-wavelet transforms has been motivated by the philosophy – which has been supported by experimental verifications – that they have the potential to combine the best of the two worlds they bridge:

1. **Wavelets:** with the power of multiresolution analysis, scaling properties of basis functions, scaling properties of wavelet coefficients, as well as the fast wavelet transform,
2. **Fractals:** with properties of scaling, recursivity and (local) self-similarity.

It is an acceptable criticism that to be considered as viable competitors in compression, hybrid methods such as FW transforms should be

required to perform *better* than the methods they hybridize, i.e., fractal coding and wavelet coding. Admittedly, this is not yet the case: To date, FW transforms do not, in general, perform better than state-of-the-art wavelet coders. For this reason, we have been examining the role of quantization and entropy coding schemes (including context-based coding) in an attempt to push fractal-wavelet coding to the limit. As well, we have been exploring the following avenues:

1. Using FW methods to improve rate-distortion performance of wavelet coders,
2. Using wavelet coder methods to improve FW coders.

Below are summarized some encouraging results of recent experiments conducted by two graduate students at Waterloo.

A. "Fractal postprocessing" (M. Ghazel). The SPIHT wavelet coder locates and transmits the most significant wavelet coefficients in terms of bit planes. It uses a hierarchical set partitioning method to identify and transmit significant wavelet coefficients while progressively varying the order of significance. For a given compression ratio, the decoder transmits the most significant bits of the most significant wavelet coefficients. All other coefficients, deemed insignificant by the encoder, are set to zero.

For various compression ratios, a fractal-wavelet transform was used at the SPIHT decoder end to interpolate between the transmitted significant coefficients in order to estimate the insignificant wavelet coefficients that were "zeroed." The most important feature is that this fractal coding is "free" since it is applied at the decoder end, requiring no fractal parameters to be sent by the SPIHT coder. A successful FW prediction will improve the fidelity of the reconstructed image without changing the bit rate.

Figure 5 illustrates the results of this experiment as applied to the 512×512 pixel, 8bpp, image *Lena*. Note that at compression ratios higher than $50:1$, the scheme actually degrades the SPIHT representation. A possible explanation is that at such high compression ratios, the SPIHT scheme has not transmitted a sufficient amount of information on the wavelet coefficients to be able to permit reliable predictions of the missing coefficients by the FW method. Setting the insignificant coefficients to zero yields better results.

At lower compression ratios, the FW scheme actually improves the quality with no "bit cost." The only costs are the extra computations required in the FW encoding/decoding, which may not be a problem if the postprocessing does not have to be performed in real time. These preliminary results show that there is potential in the use of *free* fractal-based methods to enhance SPIHT or other bit-plane coders.

B. "Fractal preprocessing" and a "Bitplane algorithm" (S. Alexander). As mentioned earlier, in zerotree/bitplaning methods such as SPIHT many small wavelet coefficients, representing fine detail in an image, are deleted in the reconstructed image. The proposed hybrid algorithm – certainly representing a different way of thinking – approximates

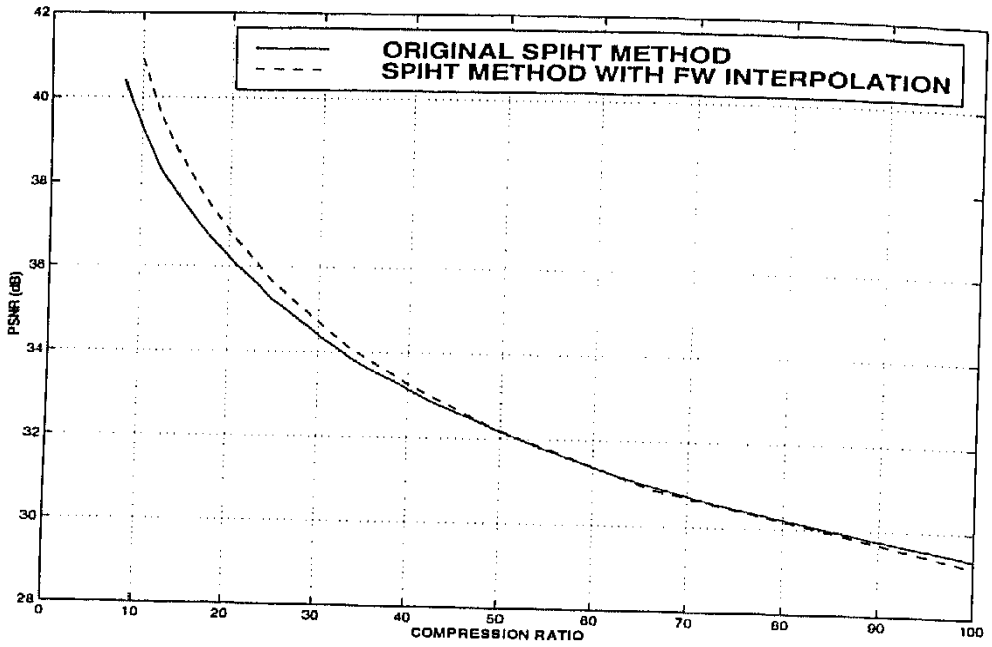

FIG. 5. *"Fractal postprocessing" of SPIHT wavelet code using FW interpolation at the decoder. (Courtesy of M. Ghazel.)*

these small coefficients using the FW coder with a low bit expense (small parent and child levels). This is the "fractal preprocessing." These coefficients are then generated and the error image (in the wavelet domain) is computed. A bitplaning algorithm is then applied to the error image.

The rationale behind this approach is to use FW coding to obtain an approximation to fine detail coefficients without too much expense and then letting the bitplane method reduce the error in this approximation. It is unlikely that such a method can produce an improvement over direct bitplaning in the $\mathcal{L}^2$ distance. Instead, the attempt is to achieve a visual improvement in compressed images without a significant cost in the rate-distortion sense. The fixed parent-child relationship is clearly wrong for most rates and should be determined based on the required rate. Nevertheless, there should be a region of the rate-distortion curve in which the fixed choice is appropriate, resulting in improvement. In order to reduce the error in the FW coder approximation selectively for the small wavelet coefficients, a modified least-squares fit is used in the collage coding which ignores larger coefficients. The gain from this approach is modest but consistent.

Another approach – the *bitplane algorithm* - has been devised to determine whether the zerotree/bitplane methods used in the SPIHT coder can be performed by a context-based arithmetic coder. Very briefly, the wavelet coefficients are first rescaled to the interval $[-1, 1]$. The algorithm begins by scanning the coefficients, evaluating significance and insignifance with a starting threshold magnitude of 0.5. If the value is significant the coefficient is coded, otherwise it is ignored. When the scan is completed

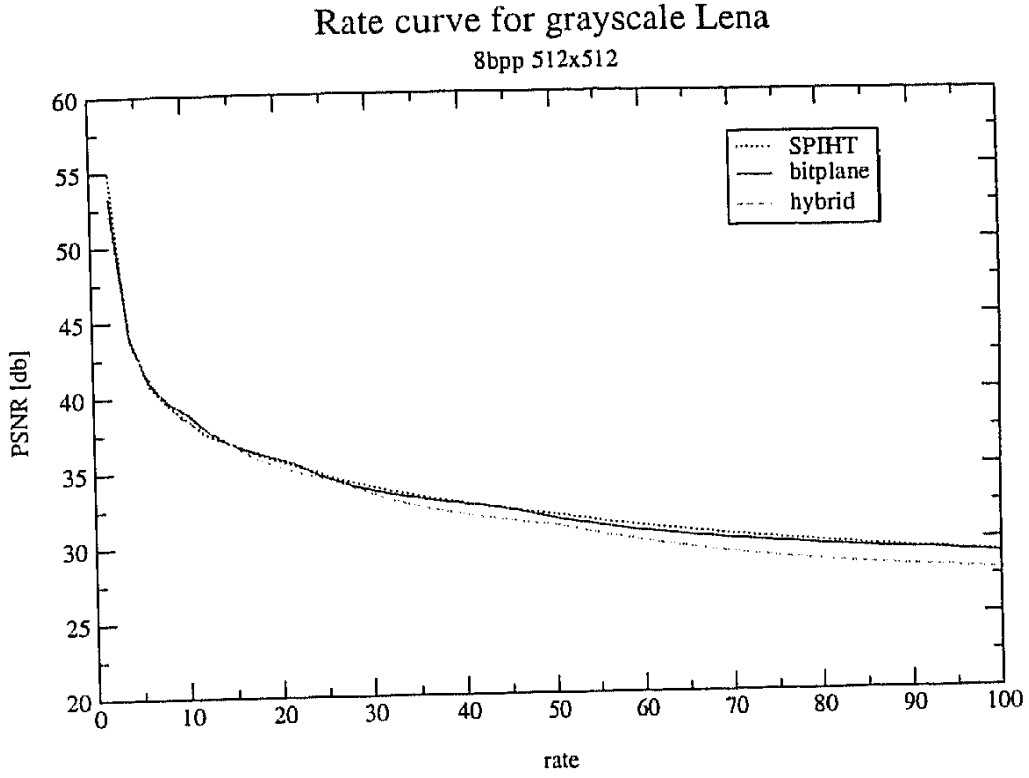

FIG. 6. *Comparison of simple bitplane, "fractal preprocessing" and SPIHT wavelet coding algorithms. (Courtesy of S.K. Alexander.)*

the threshold is reduced by one-half and the scanning is repeated. Now, however, if a coefficient was previously determined to be significant, the error is improved by a fixed amount in the appropriate direction (at a cost of 1 bit). If newly significant, the coefficient is coded as before. This process is continued until the bit budget is exhausted. The key feature is that the modelling is performed entirely in terms of the context generated by the significance of coefficients related to the current position (e.g. parent, neighbour, etc.).

Figure 6 shows the results of applying these two algorithms as well as the SPIHT coder to the same *Lena* image as above. These algorithms have been applied to many other images with quite similar results. Note that the performance of the bitplane algorithm is quite comparable to that of SPIHT. The hybrid "fractal preprocessing" algorithm also performs quite comparably, losing ground at higher rates, however. There do not seem to be any notable benefits to this method - both visually and in the $\mathcal{L}^2$ sense - when compared with the bitplane method.

4.3.3. Image analysis using fractal-wavelet transforms. The "fractal" in fractal image coding refers, of course, to the fact that an image is being aproximated by the attractor of an IFS-type operator, such attractors being typically fractal in nature (in the limit of infinite resolution and iteration). Unfortunately, this is where the "fractal analysis" of FIC stops, apart from some possible estimates of fractal dimensions and the

like. However, the concept of fractal dimension alone has little to offer to image analysis, as has been acknowledged to be the case in other areas of application. On the other hand, the rich subject of *multifractal analysis* [19, 21] has much to offer. Indeed, there has been much work ([1, 38] and others) showing how multifractal properties of functions can be understood in terms of their wavelet expansions, in particular the scaling properties of the latter. This indicates that fractal-wavelet transforms could provide a natural bridge between multifractal analysis and IFS-based fractal coding.

We must mention here that J. Lévy Vehel and his "Groupe fractales" at INRIA, Rocquencourt have been responsible for some of most detailed investigations of fractal and multifractal methods in in signal and image processing [42, 43]. Over the years, this group has developed very effective methods of segmentation, texture analysis, denoising, approximation and compression – see, for example, [14, 35, 44]. Of particular relevance to the discussion below is a *generalized IFS* (GIFS) method [14] that produces fractal interpolation functions with prescribed local regularity, in terms of their Lipschitz-Hölder exponents $\alpha(x)$.

We now show how the relationship between regularity of functions and the scaling properties of their wavelet expansions can be connected to the fractal-wavelet transform. First, recall that the FW transform performs a kind of extrapolation of wavelet coefficients onto the child wavelet blocks. Because a given parent subtree can contain several child subtrees, there is much mixing in the copying. Nevertheless, we can make some crude estimations of the asymptotic properties of the extrapolated wavelet coefficients. For simplicity, we consider the one-dimensional case and examine the following FW transform,

$$M : C_{00} \rightarrow \begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{c_{00}} \\ \hline \alpha_1 C_{00} & \alpha_2 C_{00} \\ \hline \end{array}\ , \tag{91}$$

which is related to the transform in Eq. (57) induced by the IFSM in Eq. (56). Iteration of M produces the following extrapolation of wavelet coefficients:

(92) c_{00}

1							
α_1				α_2			
α_1^2		$\alpha_1\alpha_2$		$\alpha_1\alpha_2$		α_2^2	
α_1^3	$\alpha_1^2\alpha_2$	$\alpha_1^2\alpha_2$	$\alpha_1\alpha_2^2$	$\alpha_1^2\alpha_2$	$\alpha_1\alpha_2^2$	$\alpha_1\alpha_2^2$	α_2^3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

In this case, the asymptotic behaviour of the coefficients in lower trees is given by

$$|c_{nj}| = O(|\alpha|^n), \qquad 0 \le j \le 2^n - 1, \tag{93}$$

where $\alpha = \max_i\{|\alpha_i|\}$, i.e., a geometric decay as $n \rightarrow \infty$.

We now employ some results that relate the regularity/irregularity of a function and its wavelet transform, keeping the discussion as brief and simple as possible and restricting it to the one-dimensional case. Let $f : \mathbf{R} \to \mathbf{R}$ be uniformly Lipschitz-β on an interval I where $n < \beta < n+1$ for some nonnegative integer n. That is, we assume that for all $x_0 \in I$,

$$|f(x) - P_{n,x_0}(x)| \leq K|x - x_0|^\beta, \tag{94}$$

where $P_{n,x_0}(x)$ denotes the Taylor polynomial of f at x_0 and K is independent of x_0. (See, for example, [48], p. 164.) Now let $\psi(x)$ be a wavelet with n vanishing moments. Then the continuous wavelet transform of $f(x)$ at the scale $s > 0$ and position $x_0 \in I$, defined by

$$Wf(x_0, s) = s^{-1/2} \int_{-\infty}^{\infty} \psi^* \left(\frac{x - x_0}{s} \right) dx \tag{95}$$

(assuming the integral exists), behaves as

$$|Wf(x_0, s)| \leq A s^{\beta + \frac{1}{2}}, \quad \forall x_0 \in I, \quad s > 0. \tag{96}$$

The dyadic wavelet expansion coefficients c_{nj} in Eq. (55) correspond to the discrete scales $s = 2^{-n}$, implying that

$$|c_{nj}| \leq B 2^{-n(\beta + \frac{1}{2})}. \tag{97}$$

The uniform Lipschitz behaviour of f on I implies a uniform asymptotic decay – more precisely, a *geometric* decay – of wavelet expansion coefficients across a resolution level n. (Analogous expressions exist for local Lipschitz behaviour about a point x_0.)

If we crudely compare this decay result with Eq. (93), then the (maximum) fractal scaling coefficient α and the Lipschitz exponent β are related as follows:

$$|\alpha| = 2^{-(\beta + \frac{1}{2})}. \tag{98}$$

A greater β value implies more regularity which, in turn, implies faster decay of the c_{nj} via smaller α scaling coefficients. At smooth regions of an image, we expect the wavelet coefficients c_{nj} to decay more quickly and the FW scaling coefficients to have smaller magnitudes. Conversely, near singularities, e.g. edges, we expect the FW scaling coefficients to have larger magnitudes. This is seen in Figure 7, in which magnitudes of the α_{ij} coefficients obtained from FW coding of the *Lena* image, $(k_1^*, k_2^*) = (5, 6)$ are plotted on a 64×64 grid. For simplicity, the Euclidean lengths of the vectors $(\alpha_{ij}^h, \alpha_{ij}^v, \alpha_{ij}^d)$ are plotted for $1 \leq i, j \leq 64$ so that the horizontal, vertical and diagonal contributions have been combined into one index. Larger values of $\| \alpha \|$, represented by darker squares, are clustered in irregular regions of the image, i.e., edges.

FIG. 7. *Magnitudes of the scaling coefficient vectors* $\alpha_{ij} = (\alpha_{ij}^h, \alpha_{ij}^v, \alpha_{ij}^d)$ *for a* $(5,6)$ *fractal-wavelet approximation of the Lena image.*

The above also suggests that a kind of "denoising" of the FW approximation to an image f can be performed by decreasing the magnitudes of FW scaling coefficients α_{ij}. In order to increase the Lipschitz exponent β in a region of the attractor by a factor $\Delta\beta > 0$, we multiply the appropriate scaling coefficient(s) by the factor $2^{-\Delta\beta} < 1$. We note the similarity of this modification to that of the "operator design" associated with multifractal image denoising [35, 44]. In that procedure, the wavelet coefficients c_{nj} are multiplied by the factor $2^{-n\Delta\beta}$.

Let us now return to Eq. (98). The Lipschitz exponent β is usually obtained by estimating the slope of appropriate log-log plots of wavelet coefficients c_{nj} across scales [48]. The FW transform estimates this scaling by "collaging", seeking to express lower wavelet trees as scaled copies of higher wavelet trees, essentially performing a geometric fit across scales.

Admittedly, the above analysis is very rudimentary. A more detailed analysis would have to take local regularity and all of its intricacies into consideration. Nevertheless, we hope that these results represent a scratching of the surface in what can possibly be accomplished via the fractal-wavelet transform. Another goal is to incorporate more aspects of multifractal analysis into the FW transform.

4.4. Fractal-wavelet transforms over nonseparable wavelet bases. We finally mention that two algorithms for fractal-wavelet transforms and compression have been developed for the case of nonseparable wavelet bases – bases that are *not* constructed as tensor products of wavelets on lower dimensional spaces [50]. One algorithm implements a pe-

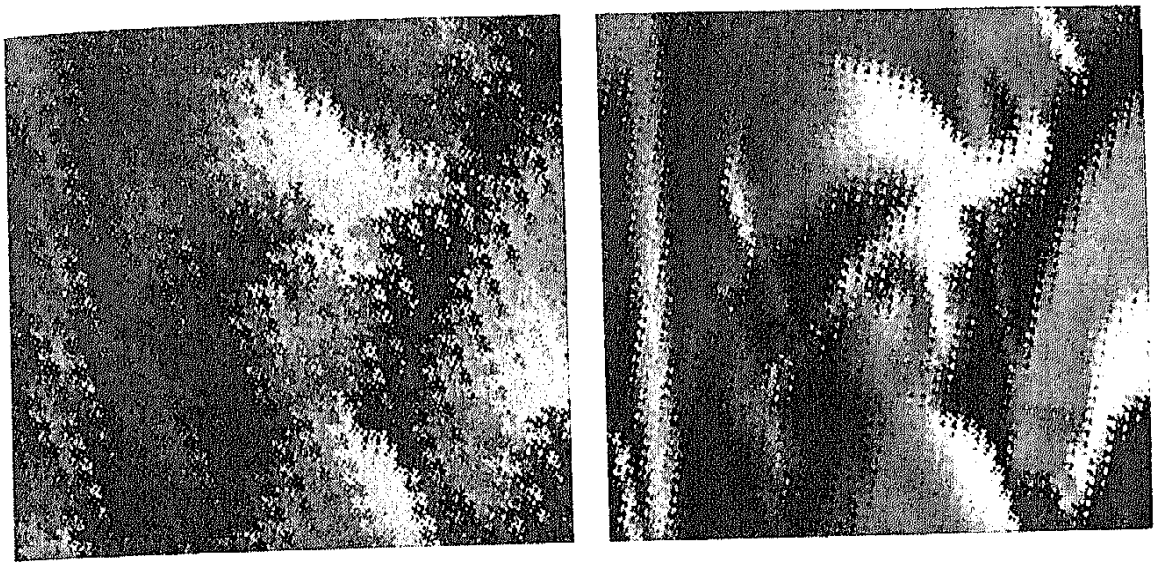

FIG. 8. *Partial wavelet expansions of the Lenna image in the Haar basis using* $2+i$ *complex tilings. Left: Summation to level* $p = 9$. *Right: Summation to level* $p = 10$. *(Courtesy of D.G. Piché.)*

riodic wavelet transform for any valid wavelet filter sequence and dilation matrices satisfying a trace condition.

The other algorithm formulates a Haar wavelet transform on tiles associated with complex bases. The characterization of multidimensional Haar wavelets was done by Gröchenig and Madych [34]. Gilbert [31] provided the connection between fractal tiles of complex bases and Iterated Function Systems. This led to a long division algorithm for complex bases [32] that provides the basis for the wavelet transform in this algorithm – essentially a translation of the Mallat decomposition algorithm into the language of complex bases.

There are cases where the two algorithms overlap – for example, the classic *twin dragon* tile. However, the dilation matrices associated with complex bases do not in general satisfy the trace condition. As a result, the two algorithms overlap but neither is a generalization of the other.

In both cases, the tilings associated with the nonseparable wavelets are nontrivial and usually fractal, even dust-like. Such tilings introduce dithering that is not concentrated along horizontal or vertical lines as is the case for separable wavelets. In Figure 8, the partial wavelet expansions of the Lenna image using the complex basis $2 + i$, for which the tiling is dust-like, are shown. The artifacts that result are visually most interesting. It is conceivable that such tilings could be useful in digital "paintbrush" applications.

5. Fractal-based methods in other areas of mathematics. The central mathematical idea behind fractal image compression is the approximation of a target element $y \in Y$, where (Y, d_Y) is an appropriate metric space, by the fixed point $\bar{y}$ of a contraction mapping $T : Y \to Y$. It is

natural to ask whether this idea can be applied to problems in other areas of mathematics that employ contraction mappings. Indeed, one of the first such areas to come to mind is that of ordinary differential equations (ODEs). If $f : \mathbf{R}^n \to \mathbf{R}^n$, then the existence and uniqueness of a solution to the initial value problem

$$\dot{x} = f(x), \qquad x(t_0) = x_0, \tag{99}$$

can be established using the associated Picard integral operator T whose action is defined by

$$(Tu)(t) = x_0 + \int_{t_0}^{t} f(u(s))ds. \tag{100}$$

With suitable conditions on f, the mapping T is contractive on an appropriate space of functions supported on an interval $[0, a]$ for some $a > 0$. This implies the existence of a unique fixed point $\bar{u} = T\bar{u}$ which is the solution to Eq. (99).

We now have the ingredients for an inverse problem: Given a target curve $y(t) \in \mathbf{R}^n$, find an ODE $\dot{x} = f(x)$ that admits $y(t)$ as either a solution or an approximate solution. The problem is to determine the optimal function f which defines the integral Picard operator. This inverse problem can be treated by the Collage Theorem [41]. Details of this approach have been given by H. Kunze in this workshop. The encouraging results obtained from this work have led us to consider inverse problems involving integral transforms, boundary-value problems and eigenvalue problems. Indeed, we are inspired to think about the possibility of using such approaches in other areas of mathematics.

Acknowledgements. As can be seen from the references, much of the earlier work was done in collaboration with Bruno Forte, whose determination to see the "big picture" of fractal transforms and fractal representation/compression of images has never faltered. Franklin Mendivil later joined the collaboration, bringing with him an energy and determination to explore a multitude of fractal-based methods both theoretically as well as practically. More recently, Herb Kunze has joined the forces, with work on the use of fractal-based methods to solve inverse problems in differential equations.

The research over the years was naturally motivated and influenced by the efforts of many "fractalators" throughout the world, a good number of whom were fortunately able to participate in this workshop. I acknowledge the enjoyable discussions that have taken place over the past years with these people. In particular, I thank Dietmar Saupe for continued discussions and for the particular collaboration reported above, much of which was done while I visited Leipzig, thanks to the support of the DFG Schwerpunktprogramm *Ergodentheorie, Analysis und Effiziente Simulation Dynamischer System* (DANSE).

And, of course, much of the exciting atmosphere at Waterloo has been due to the dedicated graduate student "fractalators," including Axel van de Walle, Jason Silver, Tom Holly, Douglas Harder and, more recently, Mohsen Ghazel, Daniel Piché and Simon Alexander. I wish to thank the latter three for providing Figures 6 to 8 as well as for discussions and details of experiments conducted during the course of their graduate work at the University of Waterloo.

Finally, I wish to thank Michael Barnsley for advice and discussions over the years, beginning during my stay as an NSERC Postdoctoral Fellow at Georgia Tech, 1984-86. I also thank him for inviting me to join him as a co-organizer of this very successful workshop.

The support of this research by grants from the Natural Sciences and Engineering Council of Canada (NSERC) is very gratefully acknowledged.

REFERENCES

[1] E. Bacry, J.F. Muzy, and A. Arnéodo, Singularity spectrum of fractal signals from wavelet analysis: exact results, J. Stat. Phys. **70**, 635–674 (1993).

[2] M. Bajraktarevic, Sur une équation fonctionelle, Glasnik Mat.-Fiz. I Astr. **12**, 201–205 (1957).

[3] S. Banach, Sur les opérations dans les ensembles abstraits et leurs applications aux équations intégrales. Fund. Math. **3**, 133–181 (1922).

[4] M.F. Barnsley, Fractal interpolation functions, Constr. Approx. **2**, 303–329 (1986).

[5] M.F. Barnsley, *Fractals Everywhere*, Academic Press, New York (1988).

[6] M.F. Barnsley and S. Demko, Iterated function systems and the global construction of fractals, Proc. Roy. Soc. London A **399**, 243–275 (1985).

[7] M.F. Barnsley, S.G. Demko, J. Elton, and J.S. Geronimo, Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities, Ann. Inst. H. Poincaré **24**, 367–394 (1988).

[8] M.F. Barnsley, V. Ervin, D. Hardin, and J. Lancaster, Solution of an inverse problem for fractals and other sets, *Proc. Nat. Acad. Sci. USA* **83**, 1975–1977 (1985).

[9] M.F. Barnsley and L.P. Hurd, *Fractal Image Compression*, A.K. Peters, Wellesley, Mass. (1993).

[10] K.U. Barthel, S. Brandau, W. Hermesmeier, and G. Heising, Zerotree wavelet coding using fractal prediction, Proc. IEEE Conf. Data Compression 1997, pp. 314–317.

[11] C.A. Cabrelli, B. Forte, U.M. Molter, and E.R. Vrscay, Iterated Fuzzy Set Systems: a new approach to the inverse problem for fractals and other sets, J. Math. Anal. Appl. **171**, 79–100 (1992).

[12] C.A. Cabrelli and U.M. Molter, Generalized self-similarity, J. Math. Anal. Appl. **230**, 251–260 (1999).

[13] P. Centore and E.R. Vrscay, Continuity properties for attractors and invariant measures for iterated function systems, Canadian Math. Bull. **37** 315–329 (1994).

[14] K. Daoudi, J. Lévy Véhel, and Y. Meyer, Construction of continuous functions with prescribed local regularity, Constr. Approx. **14**, 349–385 (1998).

[15] I. Daubechies, *Ten Lectures on Wavelets*, SIAM Press, Philadelphia (1992).

[16] G. Davis, A wavelet-based analysis fractal image compression, IEEE Trans. Image Proc. **7**, 141–154 (1998).

[17] P. Diamond and P. Kloeden, Metric spaces of fuzzy sets, Fuzzy Sets and Systems **35**, 241–249 (1990).

[18] S. Dubuc, Interpolation fractale, in *Fractal Geometry and Analysis*, J. Béelair and S. Dubuc, Eds., NATO ASI Series C, Vol. 346, Kluwer, Dordrecht (1991).

[19] C.J.G. Evertesz and B.B. Mandelbrot, Multifractal measures, in *Chaos and Fractals: New Frontiers of Science*, H.-O. Peitgen, H. Jürgens and D. Saupe, Springer Verlag, New York (1994).

[20] K. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, Cambridge (1985).

[21] K. Falconer, *Techniques in Fractal Geometry*, Wiley, Chichester (1997).

[22] Y. Fisher, A discussion of fractal image compression, in *Chaos and Fractals, New Frontiers of Science*, H.-O. Peitgen, H. Jürgens, and D. Saupe, Springer-Verlag, Heidelberg (1994).

[23] Y. Fisher, *Fractal Image Compression, Theory and Application*, Springer-Verlag, New York (1995).

[24] B. Forte, M. LoSchiavo, and E.R. Vrscay, Continuity properties of attractors for iterated fuzzy set systems, J. Aust. Math. Soc. B **36**, 175–193 (1994).

[25] B. Forte, F. Mendivil, and E.R. Vrscay, "IFS-Type Operators on Integral Transforms," in *Fractals: Theory and Applications in Engineering*, ed. M. Dekking, J. Levy-Vehel, E. Lutton, and C. Tricot, Springer Verlag, London (1999).

[26] B. Forte and E.R. Vrscay, Solving the inverse problem for measures using iterated function systems: A new approach, *Adv. Appl. Prob.* **27**, 800–820 (1995).

[27] B. Forte and E.R. Vrscay, Solving the inverse problem for functions and image approximation using iterated function systems, *Dyn. Cont. Impul. Sys.* **1** 177–231 (1995).

[28] B. Forte and E.R. Vrscay, Theory of generalized fractal transforms, in *Fractal Image Encoding and Analysis*, Y. Fisher, Ed., NATO ASI Series F 159, Springer Verlag, New York (1998).

[29] B. Forte and E.R. Vrscay, Inverse Problem Methods for Generalized Fractal Transforms, in *Fractal Image Encoding and Analysis*, *ibid.*.

[30] M. Ghazel and E.R. Vrscay, An effective hybrid fractal-wavelet image coder using quadtree partitioning and pruning, Proc. Can. Conf. Elect. Comp. Eng., CCECE 2000, Halifax, Nova Scotia (2000).

[31] W. Gilbert, Radix representations of quadratic fields, J. Math. Anal. Appl. **83**, 264–274 (1981); Fractal geometry derived from complex bases, Math. Intelligencer, **4**, 78–86 (1981); Geometry of radix expansions, in *The Geometric Vein, The Coxeter Festschrift*, C. Davis, B. Grünbaum and F.A. Sherk, Eds., Springer Verlag, New York (1982).

[32] W. Gilbert, The division algorithm in complex bases, Can. Math. Bull. **39**, 47–54 (1996).

[33] M. Giona, Vector analysis on fractal curves, in *Fractals: Theory and Applications in Engineering*, ed. M. Dekking, J. Levy-Vehel, E. Lutton, and C. Tricot, Springer Verlag, London (1999). pp. 307–323.

[34] K. Gröchenig and W.R. Madych, Multiresolution analysis, Haar bases and self-similar tilings of $\mathbf{R}^n$, IEEE Trans. Inform. Theory, **39**, 556–568 (1992).

[35] B. Guiheneuf and J. Lévy Véhel, 2-Microlocal analysis and applications in signal processing (preprint, INRIA Rocquencourt, 1997).

[36] J. Hutchinson, Fractals and self-similarity, Indiana Univ. J. Math. **30**, 713–747 (1981).

[37] A. Jacquin, *Image coding based on a fractal theory of iterated contractive image transformations*, IEEE Trans. Image Proc. **1**, 18–30 (1992).

[38] S. Jaffard, Multifractal formalism for functions, I, SIAM J. Math. Anal. **28**, 944–970 (1997).

[39] S. Karlin, Some random walks arising in learning models, I., Pacific J. Math. **3**, 725–756 (1953).

[40] H. Krupnik, D. Malah and E. Karnin, Fractal representation of images via the discrete wavelet transform, Proc. IEEE 18th Conference on Electrical Engineering (Tel-Aviv, 7–8 March 1995).

[41] H.E. Kunze and E.R. Vrscay, Solving inverse problems for ordinary differential equations using the Picard contraction mapping, Inverse Problems, **15**, 745–770 (1999).
[42] J. Lévy Véhel, Fractal approaches in signal processing (preprint).
[43] J. Lévy Véhel, Introduction to the multifractal analysis of images, in *Fractal Image Encoding and Analysis*, Y. Fisher, Ed., NATO ASI Series F 159, Springer Verlag, New York (1998).
[44] J. Lévy Véhel and B. Guiheneuf, Multifractal image denoising (preprint, INRIA Rocquencourt, 1997).
[45] J. Li and C.-C. Jay Kuo, Fractal wavelet coding using a rate-distortion constraint, Proc. ICIP-96, IEEE International Conference on Image Processing, Lausanne, Sept. 1996.
[46] N. Lu, *Fractal Imaging*, Academic Press, NY (1997).
[47] S.G. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, IEEE Trans. PAMI **11**(7), 674–693 (1989).
[48] S. Mallat, *A Wavelet Tour of Signal Processing*, Second Edition, Academic Press, New York (2001).
[49] P. Massopust, *Fractal Functions, Fractal Surfaces and Wavelets*, Academic Press, New York (1994).
[50] F. Mendivil and D. Piché, Two algorithms for nonseparable wavelet transforms and applications to image compression, in *Fractals: Theory and Applications in Engineering*, ed. M. Dekking, J. Levy-Vehel, E. Lutton, and C. Tricot, Springer Verlag, London (1999).
[51] F. Mendivil and E.R. Vrscay, Correspondence between fractal-wavelet transforms and Iterated Function Systems with Grey Level Maps, in *Fractals in Engineering: From Theory to Industrial Applications*, ed. J. Levy-Vehel, E. Lutton and C. Tricot, Springer Verlag, London, pp. 54–64. (1997).
[52] F. Mendivil and E.R. Vrscay, Fractal vector measures and vector calculus on planar fractal domains (preprint, 2001).
[53] D.M. Monro, A hybrid fractal transform, Proc. ICASSP **5**, 162–172 (1993).
[54] D.M. Monro and F. Dudbridge, Fractal Block Coding of Images, Electron. Lett. **28**, 1053–1054 (1992).
[55] S. Nadler, Multi-valued contraction mappings, Pacific J. Math. **30**, 475–488 (1969).
[56] A.H. Read, The solution of a functional equation, Proc. Roy. Soc. Edin. A **63**, 336–345 (1951–1952).
[57] M. Ruhl and H. Hartenstein, Optimal fractal coding is NP-hard, Proceedings of the IEEE Data Compression Conference, J. Storer and M. Cohn, Eds., Snowbird, Utah 1997.
[58] A. Said and W.P. Pearlman, A new fast and efficient image codec based on set partitioning in hierarchical trees, IEEE Trans. Circuits and Systems for Video Tech. **6**, 243–250 (1996).
[59] B. Simon, Explicit link between local fractal transform and multiresolution transform, Proc. ICIP-95, IEEE International Conference on Image Processing, Washington D.C., Oct. 1995.
[60] D.R. Smart, *Fixed Point Theorems*, Cambridge University Press, London (1974). p. 3.
[61] R. Strichartz, *A Guide to Distribution Theory and Fourier Transforms*, CRC Press, Boca Raton (1994).
[62] A. van de Walle, *Relating fractal compression to transform methods*, Master of Mathematics Thesis, Department of Applied Mathematics, University of Waterloo (1995).
[63] E.R. Vrscay, Iterated function systems: theory, applications and the inverse problem, in *Fractal Geometry and Analysis*, J. Bélair and S. Dubuc, Eds., NATO ASI Series C, Vol. 346, Kluwer, Dordrecht (1991).

[64] E.R. Vrscay, A Generalized Class of Fractal-Wavelet Transforms for Image Representation and Compression, Can. J. Elect. Comp. Eng. **23**(1–2), 69–84 (1998).
[65] E.R. Vrscay and D. Saupe, "Can one break the 'collage barrier' in fractal image coding?" in *Fractals: Theory and Applications in Engineering*, ed. M. Dekking, J. Levy-Vehel, E. Lutton, and C. Tricot, Springer Verlag, London, (1999). pp. 307–323.
[66] R.A. Wannamaker and E.R. Vrscay, Fractal Wavelet Compression of Audio Signals, J. Audio Eng. Soc. **45**(7–8), 540–553 (1997).
[67] R.F. Williams, Composition of contractions, Bol. Soc. Brasil. Mat. **2**, 55–59 (1971).
[68] S.J. Woolley and D.M. Monro, Rate/distortion performance of fractal transforms for image compression, Fractals **2**, 395–398 (1994).