Generalized fractal transforms, contractive mappings and associated inverse problems

Edward R. Vrscay

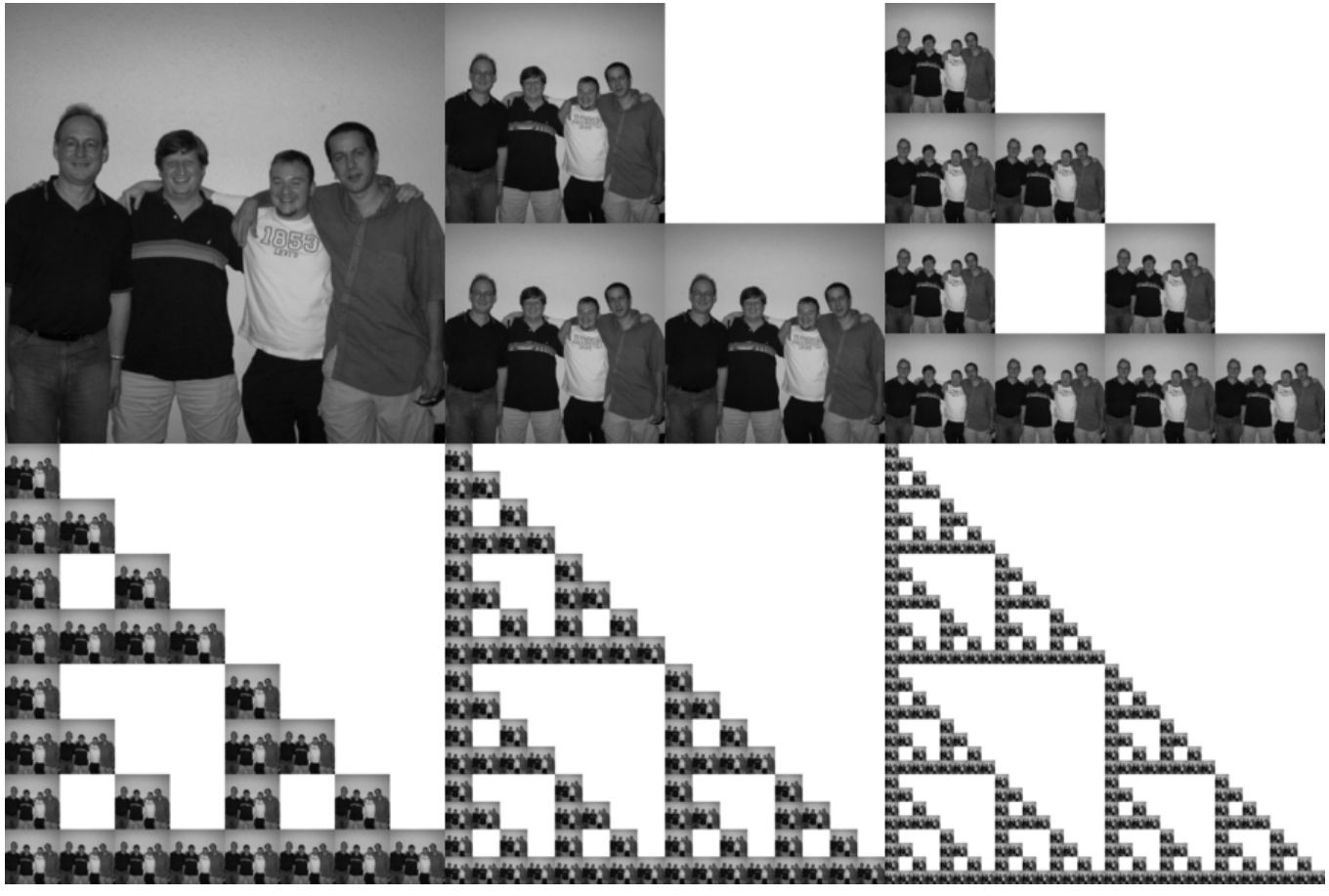
Department of Applied Mathematics, Faculty of Mathematics,
University of Waterloo, Waterloo, ON, Canada

ervrscay@uwaterloo.ca

Fractal Geometry Seminar
Wednesday, January 27, 2016
2:30 p.m. MC 5479

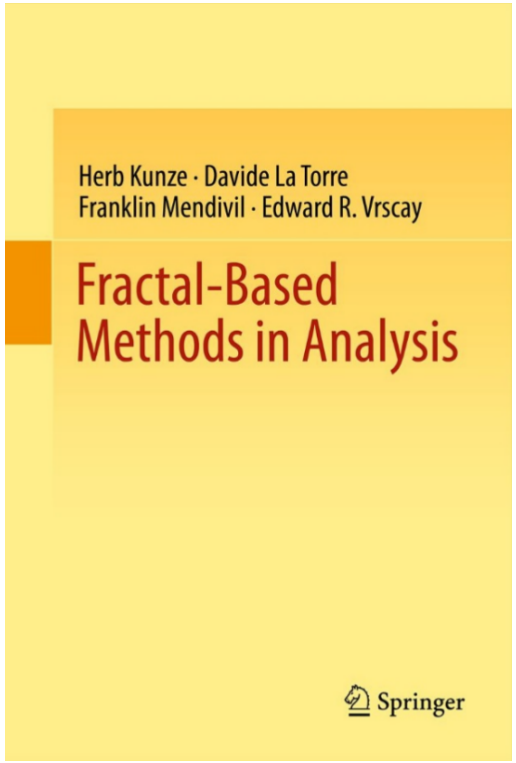
“Waterloo Fractal Analysis and Coding Project”

http://www.links.uwaterloo.ca



ERV, Herb Kunze (Guelph), Davide La Torre (Milan & Abu Dhabi), Franklin Mendivil (Acadia), ERV, Herb Kunze (Guelph), Davide La Torre (Milan & Abu Dhabi), Franklin Mendivil (Acadia), ERV, Herb Kunze (Guelph), Davide La Torre (Milan & Abu Dhabi), Franklin Mendivil (Acadia), ...

"Waterloo Fractal Analysis and Coding Project"



In memoriam:

Bruno Forte
Department of Applied Mathematics, UW (to 1993)
Universita Degli Studi de Verona, Verona, Italia (1993-2002)

Iterated Function Systems, or "Map Bags" (B. Mandelbrot)

A collection of contraction maps that operate in a parallel fashion.

This idea - in some way, shape or form - was around for quite some time, e.g.

- R.F. Williams, Composition of contractions, Bol. Soc. Brasil Mat. **2**, 55-59 (1971). Fixed points of finite compositions of contraction maps.
- S. Nadler, Multi-valued contraction mappings, Pacific J. Math. **30**, 475-488 (1969). Systems of contraction maps considered as defining "multifunctions."
- S. Karlin, Some random walks arising in learning models, I, Pacific J. Math. **3**, 725-756 (1953). Random walks over Cantor-like sets on $[0, 1]$ and associated measures - essentially the "Chaos Game" of Barnsley-Demko.

Two seminal works:

- J. Hutchinson, Fractals and self-similarity, Indiana Univ. J. Math. **30**, 713-747 (1981). Geometric and measure theoretic aspects of systems of contractive maps with associated probabilities, incl. invariant sets and probability measures supported on these sets.
- M.F. Barnsley and S.G. Demko, Iterated function systems and the global construction of fractals, Proc. Roy. Soc. London A **399**, 243-275 (1988). An independent discovery of such systems of mappings and associated attractors and invariant measures, but in a more probabilistic setting, i.e., random process. Perhaps the first solution of an "inverse problem" of fractal construction – and they tried to "match moments".

And then:

- 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). The use of the "Collage Theorem" to address the inverse problem.
- A. Jacquin, Image coding based on a fractal theory of iterated contractive image transformations, IEEE Trans. Image Proc. **1**, 18-30 (1992). Perhaps the first paper on fractal image coding.

Back to Iterated Function Systems (IFS):

Ingredients:

- (X, d): A complete metric space (e.g., $[0,1]^n$ with Euclidean metric)
- $(\mathcal{H}(X), h)$: Complete metric space of non-empty compact subsets of X with Hausdorff metric h
- $w_i : X \to X$, $1 \leq i \leq N$: Set of contraction maps on X with contraction factors $c_i \in [0,1)$.

Associated with each w_i is a set-valued mapping $\hat{w}_i : \mathcal{H}(X) \to \mathcal{H}(X)$, where

$$\hat{w}_i(S) = \{w_i(x),\, x \in S\} \quad \forall\ S \in \mathcal{H}(X)\,. \tag{1}$$

IFS operator $\hat{\mathbf{w}}$ associated with N-map IFS $\mathbf{w}$ defined as follows:

$$\hat{\mathbf{w}}(S) = \bigcup_{i=1} \hat{w}_i(S)\,, \quad S \in \mathcal{H}(X)\,. \tag{2}$$

Theorem (Hutchinson): $\hat{\mathbf{w}} : \mathcal{H}(X) \to \mathcal{H}(X)$ is contractive,

$$h(\hat{\mathbf{w}}(A), \hat{\mathbf{w}}(B)) \leq c\, h(A, B) \quad \forall\ A, B \in \mathcal{H}(X)\,, \tag{3}$$

where

$$c = \max_{1 \leq i \leq N} c_i < 1\,. \tag{4}$$

Important consequence:

From Banach's Fixed Point Theorem, there exists a unique compact set $A \in \mathcal{H}(X)$ which is the fixed point of $\hat{\mathbf{w}}$, i.e., $\hat{\mathbf{w}}(A) = A$, i.e.,

$$A = \bigcup_{i=1} A_i \quad \text{where} \quad \hat{w}_i(A)\,, 1 \leq i \leq N. \tag{5}$$

A is "self-similar," i.e., a union of contracted copies of itself.

Furthermore: For any $S_0 \in \mathcal{H}(X)$, define the iteration sequence

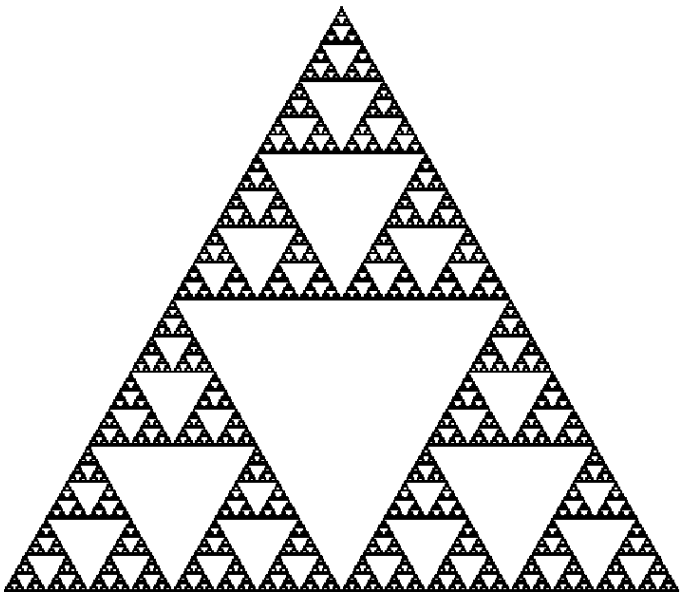
$$S_{n+1} = \hat{w}(S_n) = \bigcup_{i=1}^{N} \hat{w}_i(S_n)\,. \tag{6}$$

Then

$$\lim_{n \to \infty} h(S_n, A) = 0\,. \tag{7}$$

A is the unique (global) *attractor* of the IFS $\hat{\mathbf{w}}$.

Celebrated example: "Sierpinski triangle (gasket)"



Attractor of a 3-map affine IFS in $\mathbb{R}^2$: $w_1(x,y) = \left(\frac{1}{2}x, \frac{1}{2}y\right)$,
$w_2(x,y) = \left(\frac{1}{2}x + \frac{1}{2}, \frac{1}{2}y\right)$, $w_3(x,y) = \left(\frac{1}{2}x + \frac{1}{4}, \frac{1}{2}y + \frac{\sqrt{3}}{4}\right)$.

Another celebrated example: Barnsley's spleenwort fern

Attractor of a 4-map affine IFS in $\mathbb{R}^2$.

Of course, this leads to the question, "Can we use IFS to generate other interesting sets? Plants? Trees? Faces? ... Anything?"

This is an inverse problem

First thoughts on how to solve such an inverse problem

Suppose we have a (bounded) set $S \subset \mathbb{R}^2$, for example, another leaf-like set: Do we just start playing around with (affine) contraction maps in the plane, perturbing them, generating attractors, etc.?

Perhaps a more clever approach: We're trying to approximate S by the attractor A of an N-map IFS, i.e.,

$$S \approx A = \bigcup_{i=1}^{N} \hat{w}_i(A) = \hat{\mathbf{w}}(A)\,. \tag{8}$$

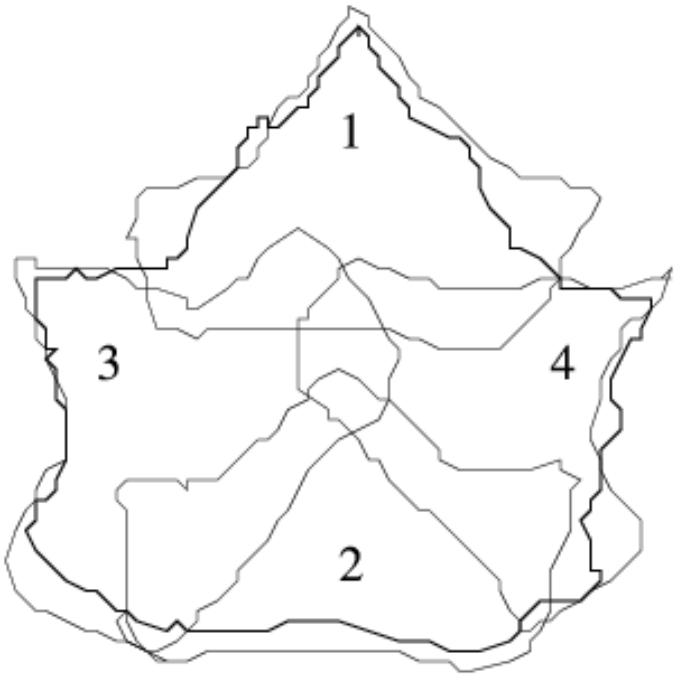
If $S \approx A$, then $\hat{w}_i(A) \approx \hat{w}_i(S)$, which implies that

$$S \approx \bigcup_{i=1}^{N} \hat{w}_i(S) = \hat{\mathbf{w}}(S)\,. \tag{9}$$

In other words: **We try to approximate S as a union of contracted copies of itself.**

Presumably, the closer S is to $\hat{\mathbf{w}}(S)$, a union , or "collage," of contracted copies of itself, the closer it is to the attractor A of the IFS $\hat{\mathbf{w}}$, i.e., the better it is approximated by A. Maybe we can express this more mathematically ... a little later.

Inverse problem by "collaging"



Left: Approximating a leaf S (darker boundary) with four contracted copies $\hat{w}_i(S)$, $1 \leq i \leq 1$ of itself. **Right:** The attractor of the resulting IFS $\hat{\mathbf{w}}$.

Banach's Fixed Point Theorem*

(Contraction Mapping Theorem ...)

is central to most (almost all?) fractal-based methods:

Let $Y, d_Y)$ be a complete metric space and $T : Y \to Y$ a contraction mapping, i.e., $d_Y(Ty_1, Ty_2) \leq c d_Y(y_1, y_2)$ for all $y_1, y_2 \in Y$ where $c \in [0, 1)$. Then there exists a unique $\bar{y} \in Y$ such that

- $T\bar{y} = \bar{y}$ (fixed point of T)
- $d_Y(T^n y_0, \bar{y}) \to 0$ as $n \to \infty$ (attractive fixed point)

*S. Banach, Sur les opérations dans les ensembles abstraites et leurs applications aux équations integrales, Fund. Math. **3** 133-181 (1922). The CMT is in an appendix to this paper, which is based on Banach's Ph.D. thesis.

Inverse problem of approximation by fixed points of contraction mappings

Let (Y, d_Y) be a complete metric space and $Con(Y)$ the set of all contraction maps $T : X \to X$. Now let $Con'(Y) \subset Con(Y)$ be a particular class of contraction maps that we wish to consider. (For example, in $\mathbb{R}^2$, the set of all N-map affine IFS, $N = 1, 2, \cdots$).

Then given a $y \in Y$ (our "target") and an an $\epsilon > 0$, can we find a $T \in Con'(Y)$ with fixed point $\bar{y}$ such that

$$d_Y(y, \bar{y}) < \epsilon \, ? \tag{10}$$

In other words, can we approximate y with the fixed point $\bar{y}$ to ϵ-accuracy?

In general, especially for fractal transforms, this problem is intractable. The following "collaging" result simplifies the problem.

The "Collage Theorem"

(That's what it's called in the fractal coding literature.)

Theorem: Let (Y, d_Y) be a complete metric space and T a contraction map on Y with contraction factor $c_T \in [0, 1)$ and fixed point $\bar{y}$. Then for any $y \in Y$,

$$d_Y(y, \bar{y}) \leq \frac{1}{1 - c_T} d_Y(Ty, y) . \tag{11}$$

$$\left[\begin{array}{c} \text{Error in approximating} \\ y \text{ with } \bar{y} \end{array} \right] \leq K(T) \left[\begin{array}{c} \text{"Collage error" in} \\ \text{approximating } Ty \text{ with } y \end{array} \right] \tag{12}$$

"Collage coding:" Try to make $d_Y(y, \bar{y})$ by finding a T that makes the collage error $d_y(y, Ty)$ as small as possible. Or rephrase as: Given a $y \in Y$ and a $\delta > 0$, find T so that

$$d_Y(Ty, y) < \delta . \tag{13}$$

Note: T does NOT have to be a fractal-type operator. More on this later.

The "Collage Theorem" was proved in an IFS setting in

- 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).

It is presented as a Remark to Banach's Theorem in

- D. Smart, *Fixed Point Theorems*, Cambridge University Press, London (1974).

Simple proof: Just play around – in the right way – with y, $\bar{y}$ and Ty, using – what else? – the triangle inequality:

$$\begin{aligned} d_Y(y,\bar{y}) &\leq d_Y(y,Ty) + d_Y(Ty,\bar{y}) \\ &= d_Y(y,Ty) + d_Y(Ty,T\bar{y}) \\ &\leq d_Y(y,Ty) + c_T d_Y(y,\bar{y})\,, \end{aligned} \tag{14}$$

and the desired result follows, i.e.,

$$d_Y(y,\bar{y}) \leq \frac{1}{1-c_T}\, d_Y(Ty,y)\,. \tag{15}$$

But, like, what if you play with y, Ty and $\bar{y}$ in the "wrong way", i.e., start with y and Ty:

$$\begin{aligned} d_Y(y, Ty) &\leq d_Y(y, \bar{y}) + d_Y(\bar{y}, Ty) \\ &= d_Y(y, \bar{y}) + d_Y(T\bar{y}, Ty) \\ &\leq d_Y(y, \bar{y}) + c_T d_Y(\bar{y}, y)\,, \end{aligned} \tag{16}$$

which yields

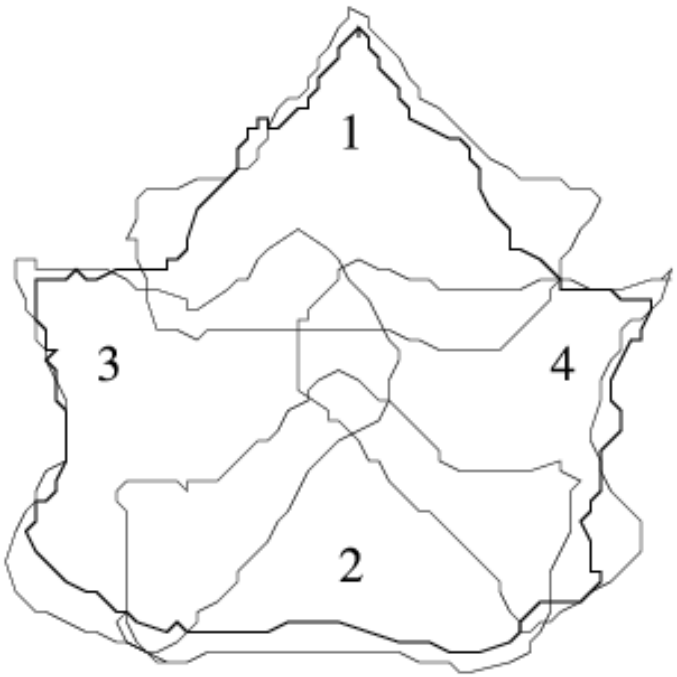
The "Anti-Collage Theorem"

$$d_Y(y, \bar{y}) \geq \frac{1}{1 + c_T} d_Y(Ty, y)\,. \tag{17}$$

Net result:

$$\frac{1}{1 + c_T} d_Y(Ty, y) \leq d_Y(y, \bar{y}) \leq \frac{1}{1 - c_T} d_Y(Ty, y)\,. \tag{18}$$

Back to the inverse problem:



If we wish to get more "realistic": Leaves – and scenes in general – are not just black-and-white. They have shading.

We have to think about having some kind of variable values associated with points/regions on the attractor sets ....

Iterated Function Systems with Probabilities (IFSP)

Ingredients:

The "IFS" part:

- (X, d): A compact metric space (e.g., $[0,1]^n$ with Euclidean metric)
- $(\mathcal{H}(X), h)$: Complete metric space of non-empty compact subsets of X with Hausdorff metric h
- $w_i : X \to X$, $1 \leq i \leq N$: Set of contraction maps on X with contraction factors $c_i \in [0,1)$.

The "P" part:

- Associated with each IFS map w_i, $1 \leq i \leq N$, is a probability $p_i \in [0,1]$, such that $\sum_{i=1}^{N} p_i = 1$.

The resulting IFSP will now operate on **measures** on X:

Let $\mathcal{M}(X)$ denote the space of probability measures on (the Borel sigma field of) X with the following (Monge-Kantorovich) metric,

$$d_M(\mu,\nu) = \sup_{f \in Lip_1(X)} \left| \int_X f \, d\mu - \int_X f \, d\nu \right| , \tag{19}$$

where

$$Lip_1(X) = \{ f : X \to \mathbb{R} , \ |f(x_1) - f(x_2)| \leq |x_1 - x_2| \ \forall x_1, x_2 \in X \} . \tag{20}$$

Theorem (Hutchinson): The metric space $\mathcal{M}(X), d_M)$ is complete.

Associated with an N-map IFSP $(\mathbf{w}, \mathbf{p})$ is an operator $M : \mathcal{M}(X) \to \mathcal{M}(X)$ defined as follows: For a $\mu \in \mathcal{M}(X)$, let $\nu = M\mu$ such that for each Borel set $S \in X$:

$$\nu(S) = (M\mu)(S) = \sum_{i=1} p_i \mu(\hat{w}_i^{-1}(S)) \,. \tag{21}$$

Theorem (Hutchinson): M is contractive in $(\mathcal{M}(X), d_M)$, i.e.,

$$d_M(M\mu, M\nu) \leq c d_M(\mu, \nu) \quad \forall \mu, \nu \in \mathcal{M}(X) \,. \tag{22}$$

(Recall that $c = \max_{1 \leq i \leq N} c_i \leq 1$.)

Corollary: There exists a unique measure $\bar{\mu} \in \mathcal{M}(X)$ such that $\mu = M\mu$. This **invariant measure** of the IFSP $(\mathbf{w}, \mathbf{p})$ satisfies the following,

$$\bar{\mu}(S) = \sum_{i=1}^{N} p_i \bar{\mu}(\hat{w}_i{}^{-1}(S)) \quad \forall S \in X \,. \tag{23}$$

In other words, μ may be expressed as a combination of spatially-contracted and translated (via the w_i) and range-altered (via the p_i) copies of itself.

Inverse problem of measure approximation using IFSP

Given a (target) measure $\mu \in \mathcal{M}(X)$, and an $\epsilon > 0$, can we find an N-map IFSP with invariant measure $\bar{\mu}$ such that

$$d_M(\mu, \bar{\mu}) < \epsilon \quad ? \tag{24}$$

Once again, we may resort to the "Collage Theorem" to consider the following inverse problem: Given a (target) measure μ and a $\delta > 0$, can we find an N-map IFSP with associated operator M so that

$$d_M(\mu, M\mu) < \delta \quad ? \tag{25}$$

Problem: It's difficult to work with measures directly. But we can work with their **moments**

"Collage Theorem for Moments:" B. Forte and ERV, Solving the inverse problem for measures using iterated function systems, Adv. Appl. Prob. **27**, 800-820 (1995).

E. Maki, Inverse problem for measures using place-dependent iterated function systems with place-dependent probabilities – M.Math. thesis (in progress) and Fractal Geometry Seminar talk, December, 2015.

But that didn't stop the "Fathers of fractal image compression":

- M.F. Barnsley and A. Sloan, A better way to compress images, BYTE Magazine, January 1988.
- A. Jacquin, Image coding based on a fractal theory of iterated contractive image transformations, IEEE Trans. Image Proc. **1**, 18-30 (1992).
- A. Jacquin, A novel fractal block-coding technique for digital images, Proc. ICASSP'90, pp. 2225-2228.

The last two papers were based on Jacquin's Ph.D. thesis in the School of Mathematics, Georgia Institute of Technology (supervised by Michael Barnsley):

- A.E. Jacquin, A fractal theory of iterated Markov operators with applications to digital image coding, Ph.D. Dissertation, Georgia Tech, 1989.

In these works, the greyscale value at a pixel of an image was treated as the measure μ of the set S represented by that pixel. The action of the operator M on these greyscale values/measures was essentially equivalent to the action of an operator that acts on functions defined on the lattice of pixels.

This naturally leads to the formulation of **fractal transforms** over function spaces.

A brief recap:

- Action of an N-map IFS operator $\hat{\mathbf{w}}$ on a set $S \in \mathcal{H}(X)$:

$$\hat{\mathbf{w}}(S) = \bigcup_{i=1}^{N} \hat{w}_i(S)\,. \tag{26}$$

- Action of the operator M associated with an N-map IFSP on a measure $\mu \in \mathcal{M}(X)$:

$$(M\mu)(S) = \sum_{i=1}^{N} p_i \mu(\hat{w}_i^{-1}(S)) \quad \forall S \in \mathcal{B}(X)\,. \tag{27}$$

Moral of the story: In both cases, an IFS-type operator "T" acts on "something" – call it "y" – by producing spatially-contracted and translated (and in the case of measures, range-modified) copies of that "something" and appropriately combining these copies to produce "something else" – call it "Ty".

This basic idea can be applied to functions, multifunctions, inclusions, multimeasures, etc.. We call such an operator T a **generalized fractal transform**.

For example, in the case of functions:

Iterated Function Systems with Greyscale Maps (IFSM)

B. Forte and ERV, Solving the inverse problem for function and image approximation using iterated function systems, Dyn. Cont. Impul. Sys. **1(2)**, 177-231 (1995).

Ingredients:

- The **base space** (or pixel space) (X, d) on which our functions will be supported: A compact metric space (e.g., $[0,1]^n$ with Euclidean metric).
- The **(image) function space** $\mathcal{F}(X) = \{u : X \to R_g\}$ where $R_g \subset \mathbb{R}$ denotes the (greyscale) range space. (In applications, $R_g \subset \mathbb{R}^+$.)
- The IFS contraction maps $w_i : X \to X$, $1 \leq i \leq N$ with contraction factors $c_i \in [0,1)$, also assumed to be one-to-one. These maps will produce the spatially-contracted and translated copies of our functions. In applications, we use affine IFS maps, e.g., $w_i(x) = s_i x + a_i$.
- The **greyscale maps**: Associated with each IFS map w_i is a greyscale map $\phi_i : R_g \to R_g$. We may also wish to consider place-dependent greyscale maps, i.e., $\phi_i : R_g \times X \to \mathbb{R}$.
  We'll assume that the ϕ_i are Lipschitz, i.e., for each $1 \leq i \leq N$, there exists a $K_i \geq 0$, such that

$$|\phi_i(t_1) - \phi_i(t_2)| \leq K_i |t_1 - t_2| \quad \forall t_1, t_2 \in R_g \,. \tag{28}$$

  In applications, we use affine ϕ-maps, i.e., $\phi_i = \alpha_i t + \beta_i$.

Associated with an N-map IFSM $(\mathbf{w}, \mathbf{\Phi})$, is the **fractal transform** $T : \mathcal{F}(X) \to \mathcal{F}(X)$ defined as follows: For a $u \in \mathcal{F}(X)$, $v = Tu$ is given by

$$v(x) = (Tu)(x) = {\sum_{i=1}^{N}}' \phi_i u((w_i^{-1}(x))) \,, \tag{29}$$

where the prime denotes summation over those $i \in \{1, \cdots, N\}$ for which $x \in w_i(X)$. In order that $v(x)$ be defined for all $x \in X$ we must have the additional condition that

$$\bigcup_{i=1}^{N} \hat{w}_i(X) = X \,, \tag{30}$$

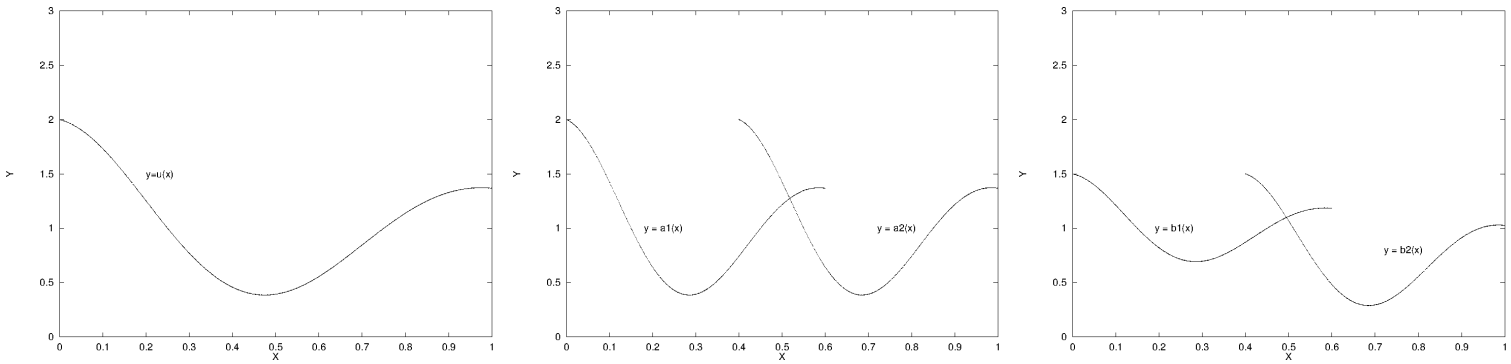
i.e., the sets $\hat{w}_i(X)$ cover X.

Action of fractal transform T on u:

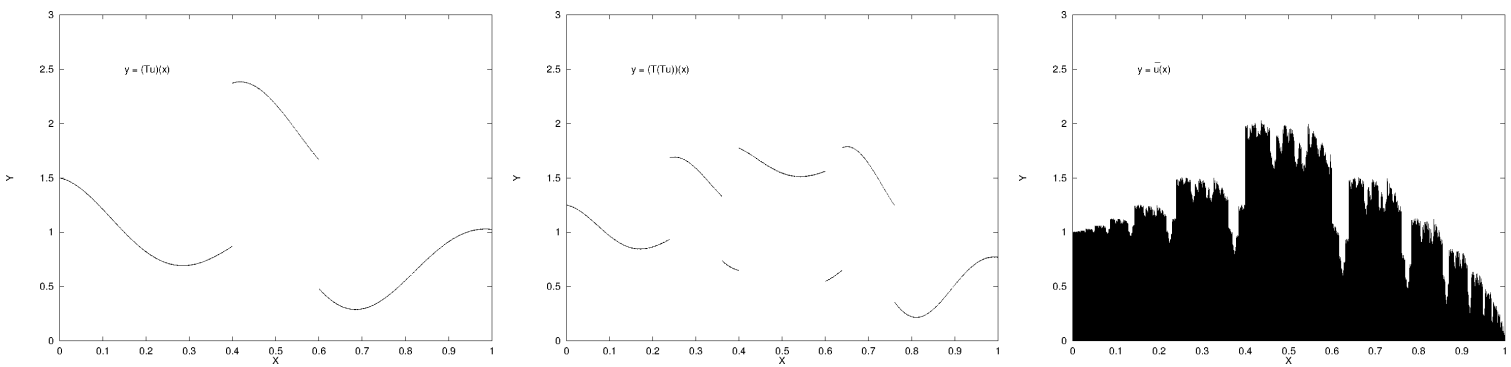
1. It produces N spatially-contracted and translated copies $u_i = u \circ w_i^{-1}$ of u.
2. It modifies the range values each of these copies: $v_i = \phi_i \circ u_i$.
3. It combines the v_i – by simple addition – to produce a new function $v = Tu$.

Example: 2-map IFSM on $X = [0, 1]$, $R_g = \mathbb{R}^+$.

$$\begin{aligned} w_1(x) &= 0.6x & \phi_1(t) &= 0.5t + 0.5 \\ w_2(x) &= 0.6x + 0.4 & \phi_2(t) &= 0.75t \end{aligned} \tag{31}$$



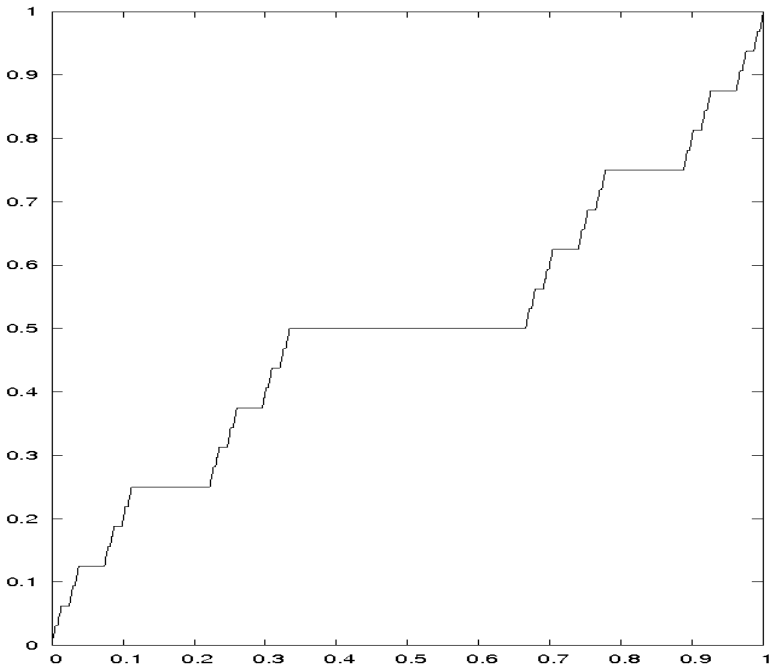
Left: The function $u(x)$. **Middle:** The spatially-contracted and translated copies $u_i(x) = u(w_i^{-1}(x))$. **Right:** The range-modified copies $v_i(x) = \phi_i(u_i(x))$.



Left: The function $v = Tu$ obtained by adding the v_i above. **Middle:** The function $w = Tv = T^2u$ **Right:** The fixed point $\bar{u} = T\bar{u}$.

Example: The following 3-map IFSM on $X = [0, 1]$, $R_g = \mathbb{R}^+$.

$$\begin{array}{ll} w_1(x) = \frac{1}{3}x & \phi_1(t) = \frac{1}{2}t \\ w_2(x) = \frac{1}{3}x + \frac{1}{3} & \phi_2(t) = \frac{1}{2} \\ w_3(x) = \frac{1}{3}x + \frac{2}{3} & \phi_2(t) = \frac{1}{2}t + \frac{1}{2} \end{array} \tag{32}$$

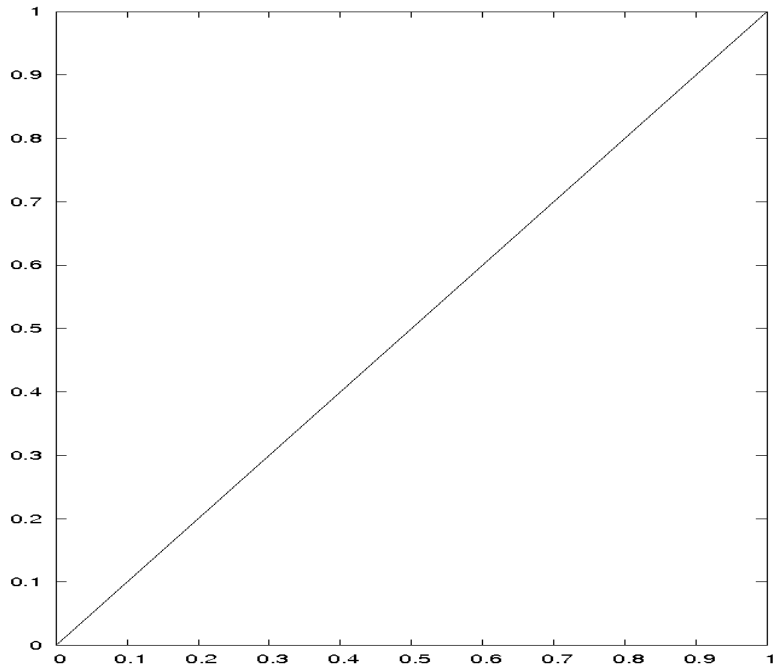


The attractor $\bar{u}$ of this IFSM: The "Devil's staircase function."

But not all fixed point functions $\bar{u}$ need to be "fractal-like"

Example: The following 2-map IFSM on $X = [0, 1]$, $R_g = \mathbb{R}^+$.

$$\begin{array}{ll} w_1(x) = \frac{1}{2}x & \phi_1(t) = \frac{1}{2}t \\ w_2(x) = \frac{1}{2}x + \frac{1}{2} & \phi_2(t) = \frac{1}{2}t + \frac{1}{2} \end{array} \tag{33}$$



The attractor $\bar{u}$ of this IFSM: $\bar{u}(x) = x$.

We haven't said anything about contractivity of IFSM operators. For $\mathcal{F}(X) = L^p(X)$,

$$
\begin{aligned}
\|Tu - Tv\|_p &\leq \left\| \sum_{i=1}^{N} \left[\phi_i \circ u \circ w_i^{-1} - \phi_i \circ v \circ w_i^{-1} \right] \right\|_p \\
&\leq \sum_{i=1}^{N} \|\phi_i \circ u \circ w_i^{-1} - \phi_i \circ v \circ w_i^{-1}\|_p \\
&\leq \sum_{i=1}^{N} \left[\int_{\hat{w}_i(X)} \left| \phi_i(u(w_i^{-1}(x))) - \phi_i(v(w_i^{-1}(x))) \right|^p \, dx \right]^{1/p} \\
&\leq \sum_{i=1}^{N} K_i \left[\int_{\hat{w}_i(X)} \left| u(w_i^{-1}(x)) - v(w_i^{-1}(x)) \right|^p \, dx \right]^{1/p} \\
&\leq \sum_{i=1}^{N} K_i c_i^{1/p} \left[\int_X |u(y) - v(y)|^p \, dy \right]^{1/p} \qquad (\, y = w_i^{-1}(x) \Longrightarrow x = w_i(y) \,) \\
&= \left[\sum_{i=1}^{N} K_i c_i^{1/p} \right] \|u - v\|_p \,. \qquad (34)
\end{aligned}
$$

Induced fractal transform operators

Let ϕ be a 1-1 mapping of our space of functions $\mathcal{F}(X)$ to a representation space $\mathcal{G}$, e.g.,

- Fourier transforms
- Orthogonal expansion (e.g., Fourier, wavelet)
- (in the case of probability measures, the moment space $(1, g_1, g_2, \cdots)$)

A fractal transform $T : \mathcal{F}(X) \longrightarrow \mathcal{F}(X)$ induces an operator $M : \mathcal{G} \longrightarrow \mathcal{G}$:

$$\begin{array}{ccccc} & & T & & \\ & u & \longrightarrow & v & \\ \phi & \downarrow & & \uparrow & \phi^{-1} \\ & U & \longrightarrow & V & \\ & & M & & \end{array} \tag{35}$$

Action of M on an element $U \in \mathcal{G}$: It produces N copies of U and recombines them to produce a $V = MU$ in $\mathcal{G}$.

Example: Fourier transforms

Let T be the fractal transform associated with an N-map affine IFSM on $\mathbb{R}$:

$$w_i(x) = s_i x + a_i \,, \qquad \phi_i(t) = \alpha_i t + \beta_i \,. \tag{36}$$

so that

$$v(x) = (Tu)(x) = \sum_{i=1}^{N}{}' \left[\alpha_i u(w_i^{-1}(x)) + \beta_i \right] \,. \tag{37}$$

Once again, $v = Tu$ is a linear combination of N **spatially-contracted**, translated and range-modified copies of v.

Operator M induced on the space of Fourier transforms:

$$V(\omega) = (MU)(\omega) = \sum_{i=1}^{N} a_k s_k e^{-a_k \omega} U(s_k \omega) + \sum_{k=1}^{N} \beta_k s_k e^{-i a_k \omega} \mathcal{F}_X(s_k \omega) \,, \tag{38}$$

where

$$\mathcal{F}_X = \int_X e^{-i\omega x} dx \,. \tag{39}$$

$V = MU$ is a linear combination of N phase-shifted, **frequency-expanded**, translated and range-modified copies of U (plus some sinc functions).

Inverse problem for IFSM

Let's go directly to "Collage coding": Given a $u \in L^p(X)$, can we find an N-map IFSM $(\mathbf{w}, \mathbf{\Phi})$ with contractive fractal transform T so that for a desired $\delta > 0$,

$$d_p(u, Tu) \leq \delta \quad ? \tag{40}$$

Usual strategy: Fix the IFS maps w_i, $1 \leq i \leq N$, and find the best ϕ_i maps, i.e., the ϕ_i maps that minimize the collage distance $d_p(u, Tu)$.

And to make life even simpler: Use "nonoverlapping" IFS maps w_i, i.e., the sets $\hat{w}_i(X)$ overlap on sets of Lebesgue measure zero.

Example: On $X = [0, 1]$,

$$w_1(x) = \frac{1}{2}x\,, \quad w_2(x) = \frac{1}{2}x + \frac{1}{2}\,. \tag{41}$$

The inverse problem "separates" into inverse problems over each each subset $X_i = \hat{w}_i(X)$. Over each subset X_i, you find the best ϕ_i map.

Inverse problem for Fourier transforms

Recall: A fractal transform $T : \mathcal{F}(X) \to \mathcal{F}(X)$ induces an operator $M : \mathcal{G} \to \mathcal{G}$:

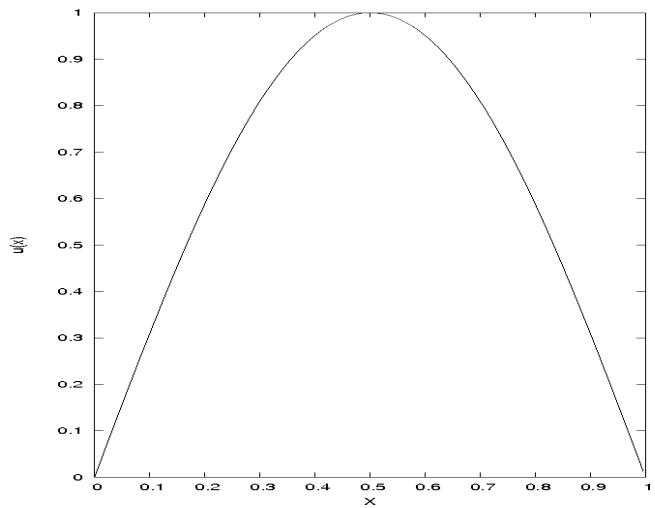
$$\begin{array}{ccccc} & & T & & \\ & u & \longrightarrow & v & \\ \phi & \downarrow & & \uparrow & \phi^{-1} \\ & U & \longrightarrow & V & \\ & & M & & \end{array} \tag{42}$$

- G. Mayer and ERV, Iterated Fourier transform systems: A method for frequency extrapolation, in *Image Analysis and Recognition*, LNCS **4633** (Springer, Berlin-Heidelberg), p. 728-739.

Motivation: The signal from an MRI is a Fourier transform $U(\omega)$. It must be inverted to produce an image $u(x)$ of the subject. Instead of working on the image u, e.g., denoising, superresolution, we work on the original signal U, and then invert.

But there's a problem: You're really pushing it if you think that functions in general can be approximated well as combinations of spatially-contracted, translated and range-modified copies of themselves.

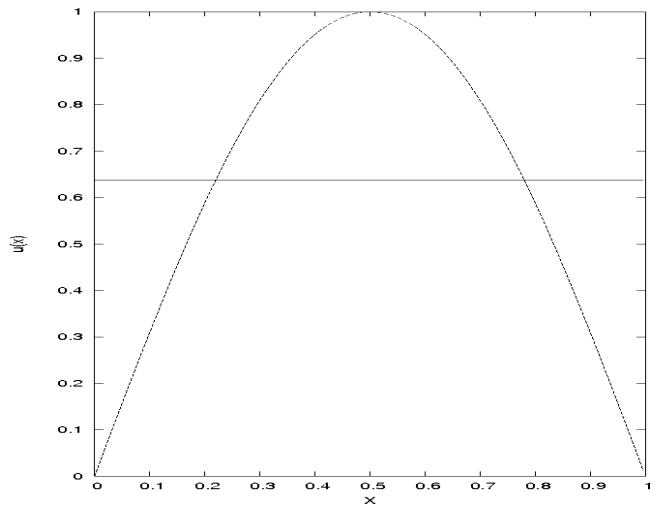
Example: Target function $u(x) = \sin \pi x$ on $[0, 1]$



Suppose $w_1(x) = \frac{1}{2}x$, $w_2(x) = \frac{1}{2}x + \frac{1}{2}$.
It doesn't look like you can approximate $u(x)$ on [0,1] with two copies of itself on $[0.1/2]$ and $[1/2, 1]$.

OK, but let's try it anyway.

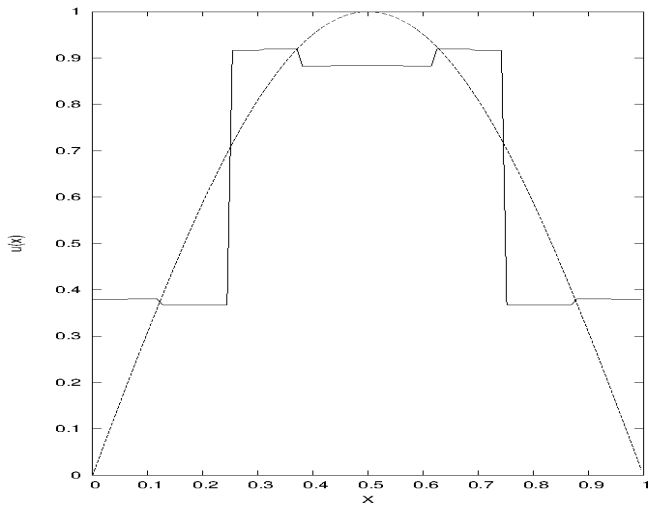
Example: Target function $u(x) = \sin \pi x$ on $[0, 1]$



The best "tile" of u with two copies of itself is the constant function $u(x) = \bar{u}$, average value of u on [0,1].

OK, let's push it to four copies of itself ...

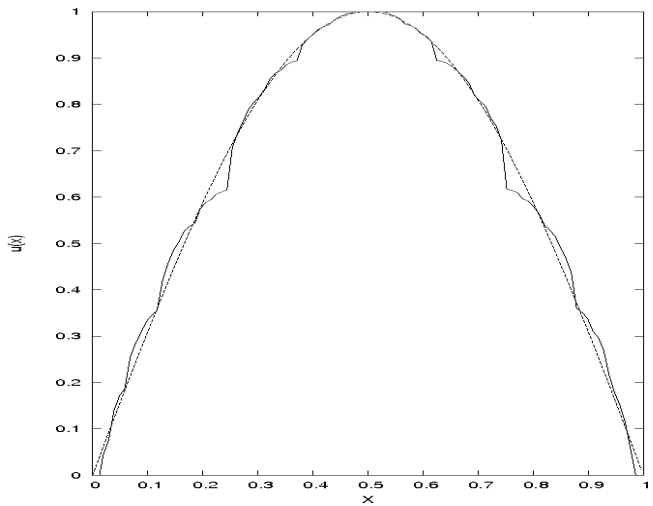
Example: Target function $u(x) = \sin \pi x$ on $[0, 1]$



The best "tile" of u with four copies of itself is a piecewise constant approximation to $u(x)$ on [0,1].

Of course, we can increase the number of tiles to produce better piecewise constant approximations to u. Rather unsavoury, n'est-ce pas?

A better strategy is to break up the function u into "pieces", i.e., consider it as a collection of functions defined over subintervals of X.



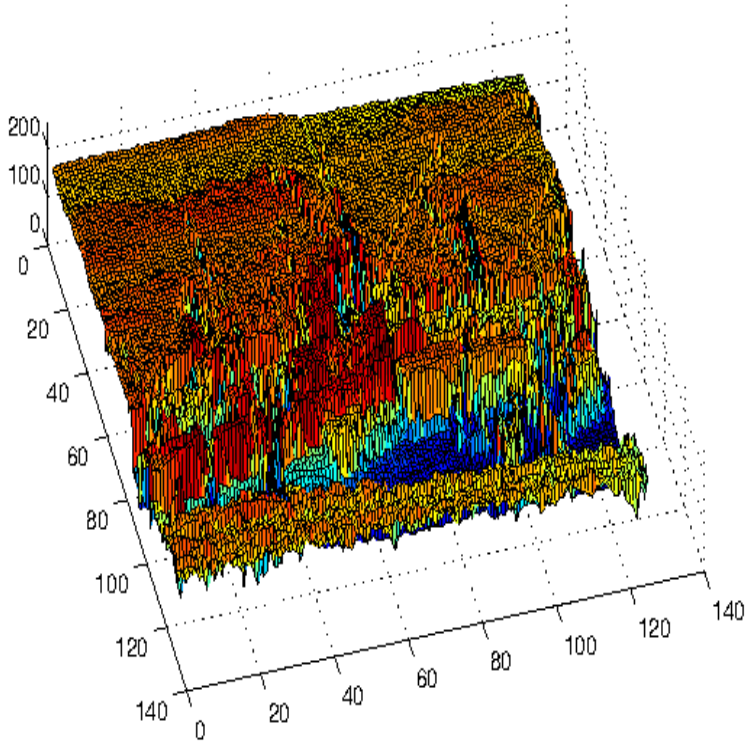
- Four "range" intervals, $[0, 1/4]$, $[1/4, 1/2]$, $[1/2, 3/4]$, $[3/4, 1]$
- Two "domain" intervals, $[0, 1/2]$, $[1/2, 1]$.

Approximate the target on each "range" interval as a spatially-contracted and range-modified copy of one of the two "pieces" of the target on the domain intervals.

A much better approximation, n'est-ce pas?

Since Jacquin (1992), such a "block-based" strategy has been used to perform **fractal image coding**.

Application to image functions $u : X \to R_g$, $X \subset \mathbb{R}^2$

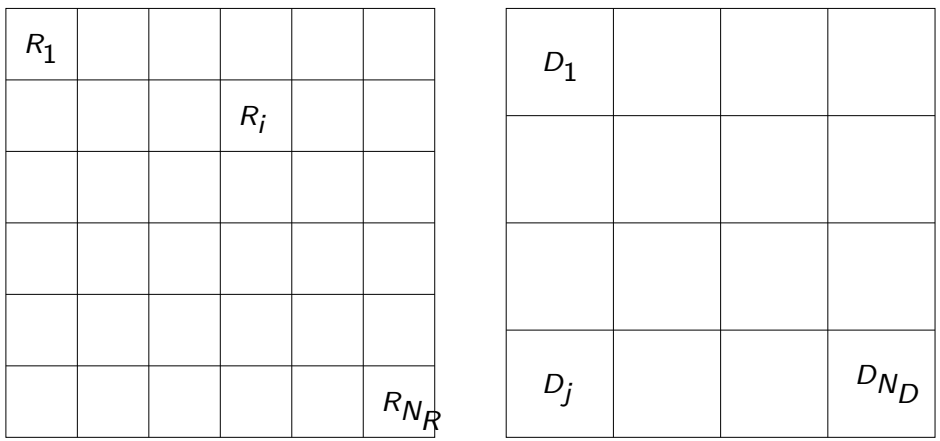


Left: The standard test-image, *Boat*, a 512×512-pixel digital image, 8 bits per pixel. Each pixel assumes one of 256 greyscale values between 0 and 255. **Right:** The *Boat* image, viewed as a non-negative image function $z = u(x, y)$. The red-blue spectrum of colours is used to characterize function values: Higher values are more red, lower values are more blue.

A simple block-based fractal coding scheme for a greyscale image function u

In what follows, let X be an $n_1 \times n_2$ pixel array on which the image u is defined.

- Let $\mathcal{R}^{(n)}$ denote a set of $n \times n$-pixel **range** subblocks R_i, $1 \leq i \leq N_{\mathcal{R}^{(n)}}$, which cover X, i.e., $X = \cup_i R_i$.
- Let $\mathcal{D}^{(m)}$ denote a set of $m \times m$-pixel **domain** D_j, $1 \leq j \leq N_{\mathcal{D}^{(m)}}$, where $m = 2n$. (The D_i are not necessarily non-overlapping, but they should cover X.)
- Let $w_{ij} : D_j \longrightarrow R_i$ denote affine geometric transformations. There are 8 such maps: 4 rotations, 2 diagonal flips, vertical and horizontal flips. In many cases, only the zero rotation map is employed. (Some kind of "decimation" is also required to map larger blocks to smaller blocks.)



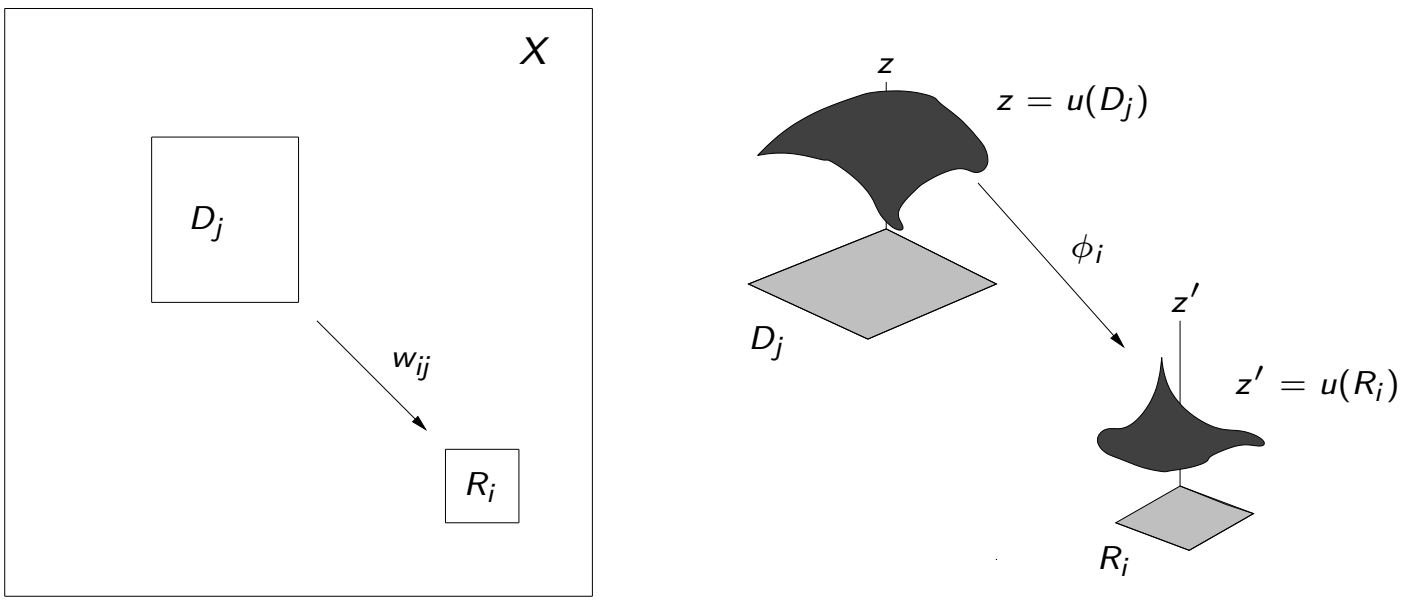
- For each range block R_i, $1 \leq i \leq N_{\mathcal{R}^{(n)}}$, compute the errors associated with the approximations

$$u(R_i) \approx \phi_{ij}(u(w_{ij}^{-1}(R_i)), \quad \text{for all} \quad 1 \leq j \leq N_{\mathcal{D}^{(m)}}\,, \tag{43}$$

where we use affine greyscale transformations,

$$\phi(t) = \alpha t + \beta\,. \tag{44}$$

Choose the domain block $j(i)$ that yields the lowest approximation error.



Left: Range block R_i and associated domain block D_j. **Right:** Greyscale mapping ϕ_i from $u(D_j)$ to $u(R_i)$.

The collection of parameters for all range blocks R_i, $1 \leq i \leq N_{\mathcal{R}^{(n)}}$,

$$\begin{array}{ll} j(i), & \text{index of best domain block,} \\ \alpha_i, \beta_i, & \text{affine greyscale map parameters,} \end{array} \tag{45}$$

comprise the **fractal code** of the image function u. The fractal code defines a fractal transform T.

The fixed point $\bar{u}$ of T is designed to approximate the image u.

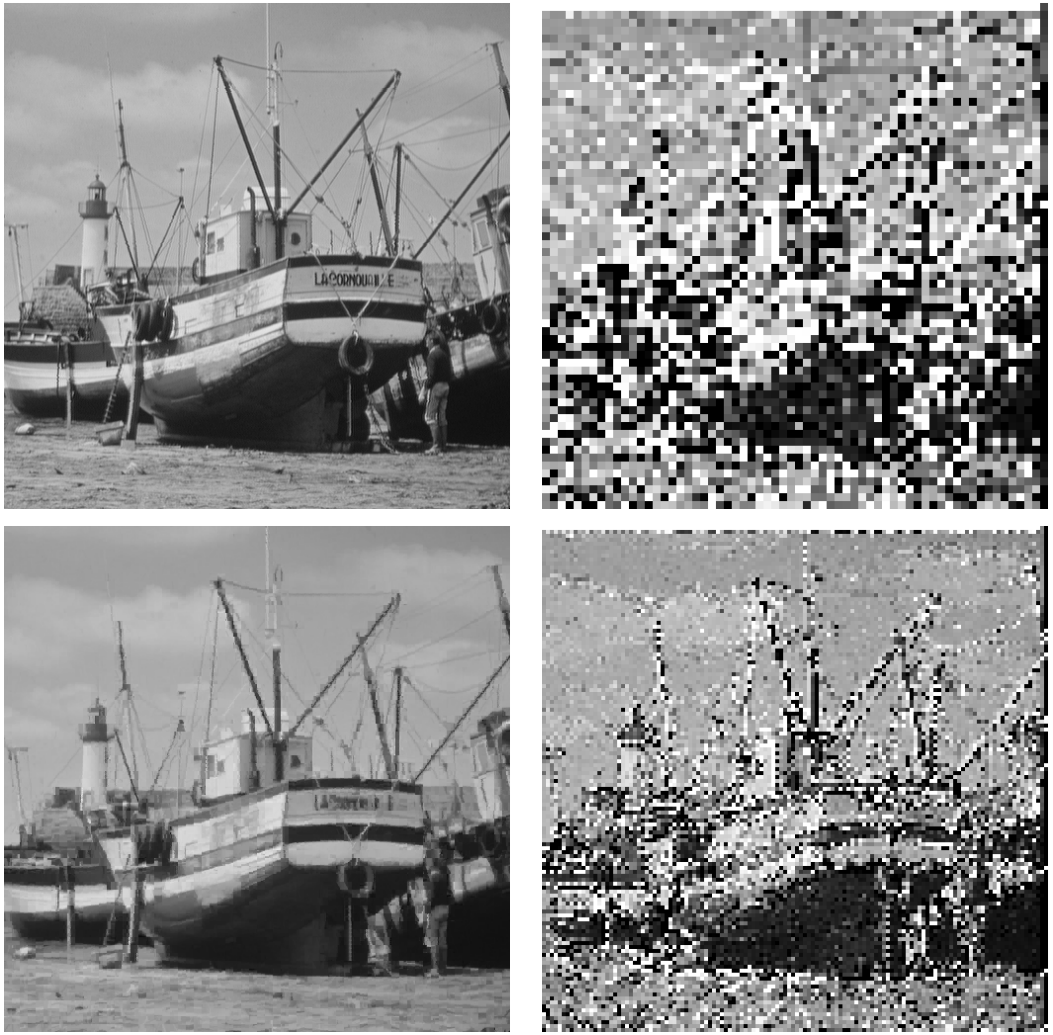
Why? Because minimization of the errors associated with the approximations

$$u(R_i) \approx \phi_{ij}(u(w_{ij}^{-1}(R_i)), \quad \text{for all} \quad 1 \leq j \leq N_{\mathcal{R}^{(n)}}\,, \tag{46}$$

is actually **minimizing the collage error Tu over the subblock $u(R_i)$**. (You're finding a part of T, T_i, that maps $u(R_i)$ as close as possible to itself.)

Moral of the story: You store the fractal code of u and generate its approximation $\bar{u}$ by iterating T, as shown in the next slide.

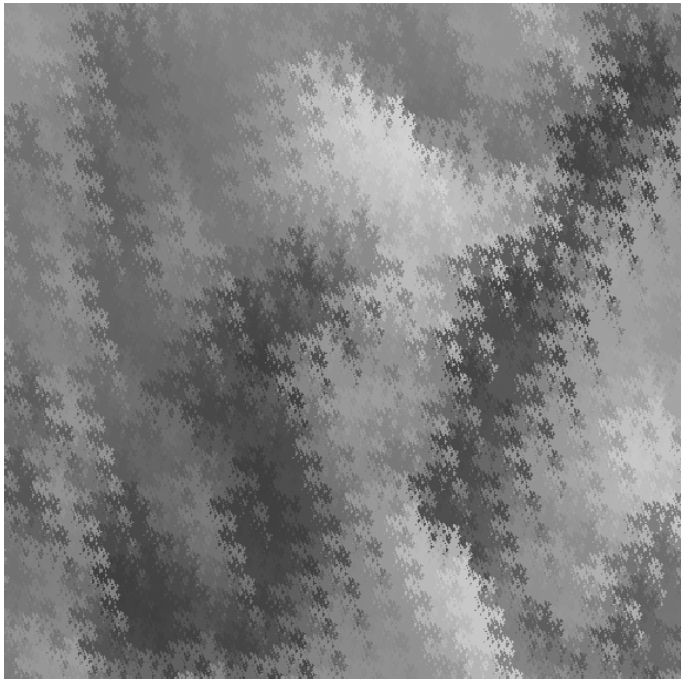
Fractal image coding



Clockwise, starting from top left: Original *Boat* image. The iterates u_1 and u_2 and fixed point approximation $\bar{u}$ obtained by iteration of fractal transform operator. ($u_0 = \mathbf{0}$.) 8×8-pixel range blocks. 16×16-pixel domain blocks.

Fractal image coding on complex tiles

(in honour of Will Gilbert)



Actually, these are the results of **fractal-wavelet coding** of the standard *Lena* image (512×512-pixel, 8 bits per pixel) using the Haar basis supported on tiles associated with the complex basis $z = 2 + i$.

Left: 9 scaling levels. **Right:** 10 scaling levels.

D.G. Piché, Complex Bases, Number Systems and Their Application to Fractal-Wavelet Image Coding, Ph.D. Thesis, Department of Applied Mathematics (2002).

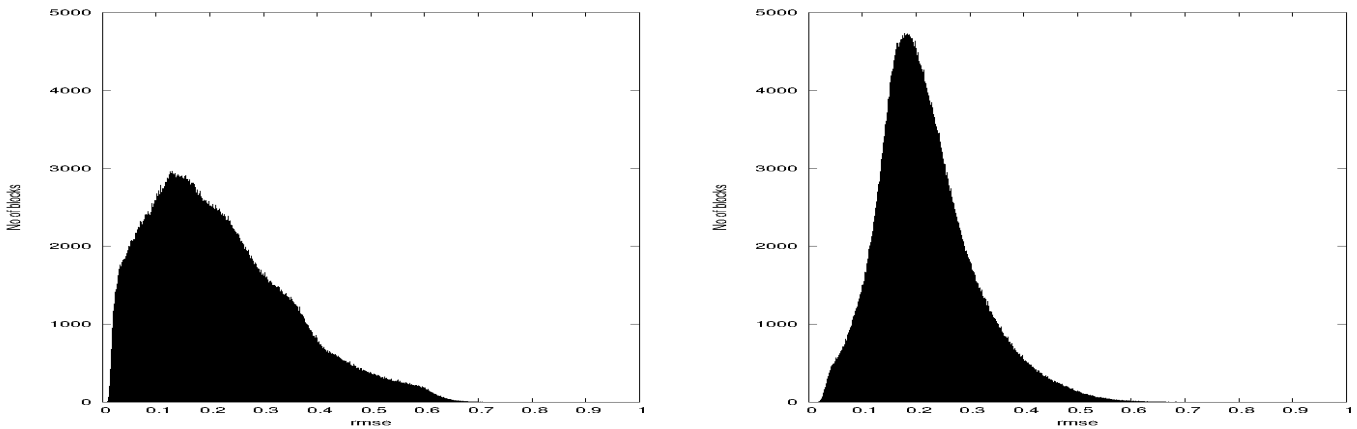
Why does fractal image coding "work"?

Because images are quite self-similar – at least locally. In

S.K. Alexander, ERV and S. Tsurumi, A simple, general model for the affine self-similarity of images, ICIAR 2008,

it was shown that images generally possess a considerable degree of local **affine self-similarity**, i.e.,

Subblocks of an image are well approximated by a number of other subblocks – with – or even without – the help of affine greyscale transformations



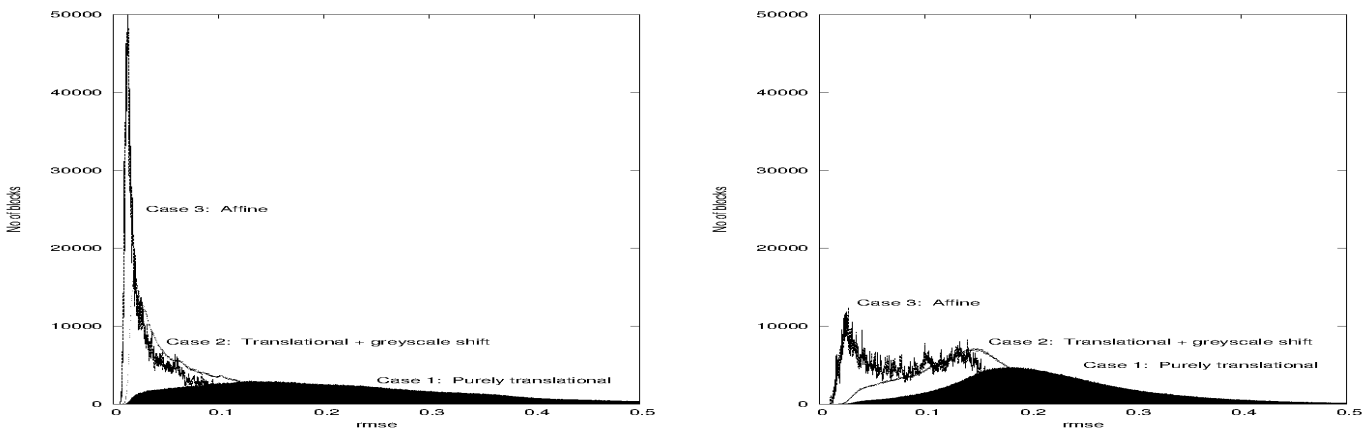
(a) Lena (b) Mandrill

Distribution of L^2 errors, $\Delta_{ij}^{(Case\ 1)} = \|u(R_j) - u(R_i)\|_2, \quad i \neq j$, in approximating all 8×8-pixel blocks $u(R_i)$ of a (normalized) image with with all other blocks $u(R_j)$.

If we allow nontrivial affine greyscale maps, e.g.

- "Case 2": $u(R_i) \approx u(R_j) + \beta \qquad (\beta = \overline{u(R_i)} - \overline{u(R_j)})$
- "Case 3": $u(R_i) \approx \alpha u(R_j) + \beta$

The error distributions are signifcantly pushed toward zero error:



(a) Lena (b) Mandrill

Same-scale (Cases 1,2 and 3) RMS self-similarity error distributions for normalized *Lena* and *Mandrill* images. Again, 8×8-pixel blocks $R_i = D_i$ were used. Case 1 distributions are shaded.

The local self-similarity of images has been implicitly used in a number of **nonlocal* image processing schemes**, including

- Nonlocal-means denoising (A. Buades, B. Coll and J.M. Morel 2005, 2010).
- Method of "self-examples," and "examples" in general (e.g., BM3D method of K. Dabov *et al.*, IEEE Trans. Image Proc. 2007).
- Fractal image coding (N. Lu, *Fractal Imaging*, Academic Press 1997).
- Yes, vector quantization! (Fractal image coding is, in fact, "self-vector quantization".)
- The list goes on and on and on ...

***Nonlocal image processing:** In constrast to many classical methods (e.g., "local averaging"), the greyscale value $u(i,j)$ at pixel (i,j) is modified according to greyscale values $u(i_k,j_k)$ at pixels (i_k,j_k) that are not necessarily in the neighbourhood of (i,j).

More recent work: Multifunction representations of images

(and associated generalized fractal transforms)

with special thanks to Davide La Torre

- **Measure-valued image functions:**
  - D. La Torre, ERV, M. Ebrahimi, M.F. Barnsley, Measure-valued images, associated fractal transforms and the affine self-similarity of images, SIAM J Imaging Sciences **2** (2), 470-507 (2009)
  - D. La Torre and ERV, Random measure-valued image functions, fractal transforms and self-simlilarity, Applied Mathematics Letters **24**, 1405-1410 (2011)
  - D. La Torre and ERV, Fourier transforms of measure-valued images, self-similarity and the inverse problem, Signal Processing **101**, 11-18 (2014).
- **Function-valued (or Banach-valued) image functions:**
  - D. Otero, D. La Torre, O. Michailovich and ERV, On function-valued mappings and their application to the processing of hyperspectral images, preprint.
  - O. Michailovich, D. La Torre and ERV, Function-valued mappings, total variation and compressed sensing for diffusion MRI, ICIAR 2012.
  - I.C. Salgado Patarroyo, S. Dolui, O.V. Michailovich and ERV, Reconstruction of HARDI data using a split Bregman optimization approach, ICIAR 2013.
  - D. La Torre, D. Otero and ERV, A simple class of fractal transforms for hyperspectral images, Applied Mathematics and Computation **231**, 435-444 (2014).
  - D. Otero, Function-valued mappings and SSIM-based optimization in imaging, Ph.D. Thesis, Dept. of Applied Mathematics, UW (2015).

Multifunction representations of images

We now generalize the usual mathematical representation of a (greyscale/colour) image, i.e.,

$$u : X \to R_g,$$

where

- X: **base** or **pixel space**, the support of the image, $X \subset \mathbb{R}^n$, $n = 1, 2, 3$,
- $R_g \subset \mathbb{R}$ (or $\mathbb{R}^3$): the **greyscale** (or **colour**) **range**.

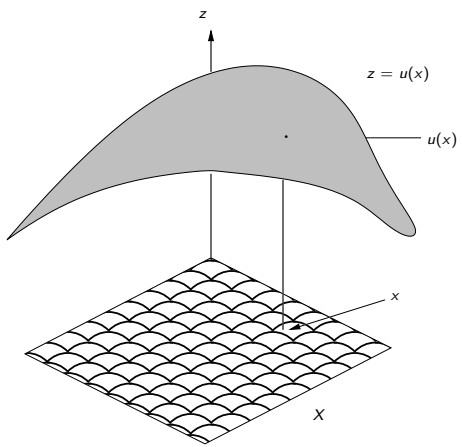
Multifunction representations of image functions I: Measure-valued

Greyscale-valued image function

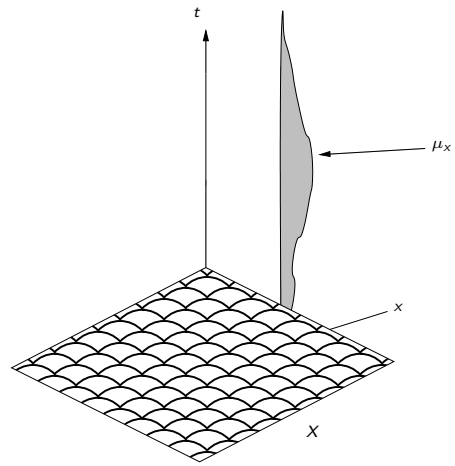
At each pixel $x \in X$, $u(x)$ is a **real value** (or a vector of real values, i.e., "RGB")

Measure-valued image function

At each pixel $x \in X$, $u(x)$ is a **measure**, i.e., μ_x



(a) Greyscale-valued image function



(b) Measure-valued image function

Example: In treatments of random images, $u(x) = \mu_x$ represents the probability distribution of the random greyscale value at $x \in X$ over the greyscale range R_g.

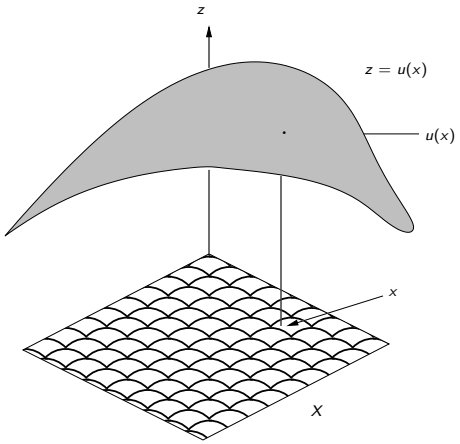
Multifunction representations of image functions II: Function-valued

Greyscale-valued image function

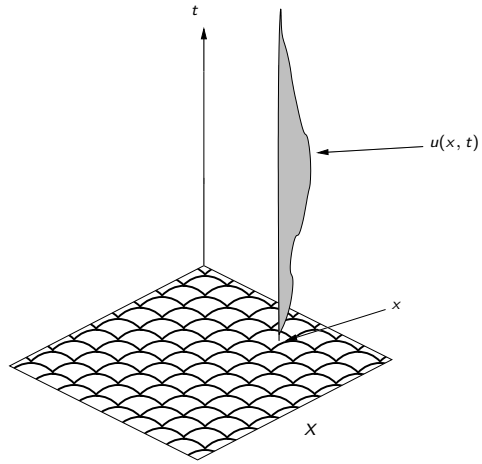
At each pixel $x \in X$, $u(x)$ is a **real value** (or a vector of real values, i.e., "RGB")

Function-valued image function

At each pixel $x \in X$, $u(x)$ is a **real-valued function**, i.e., $u(x; t)$



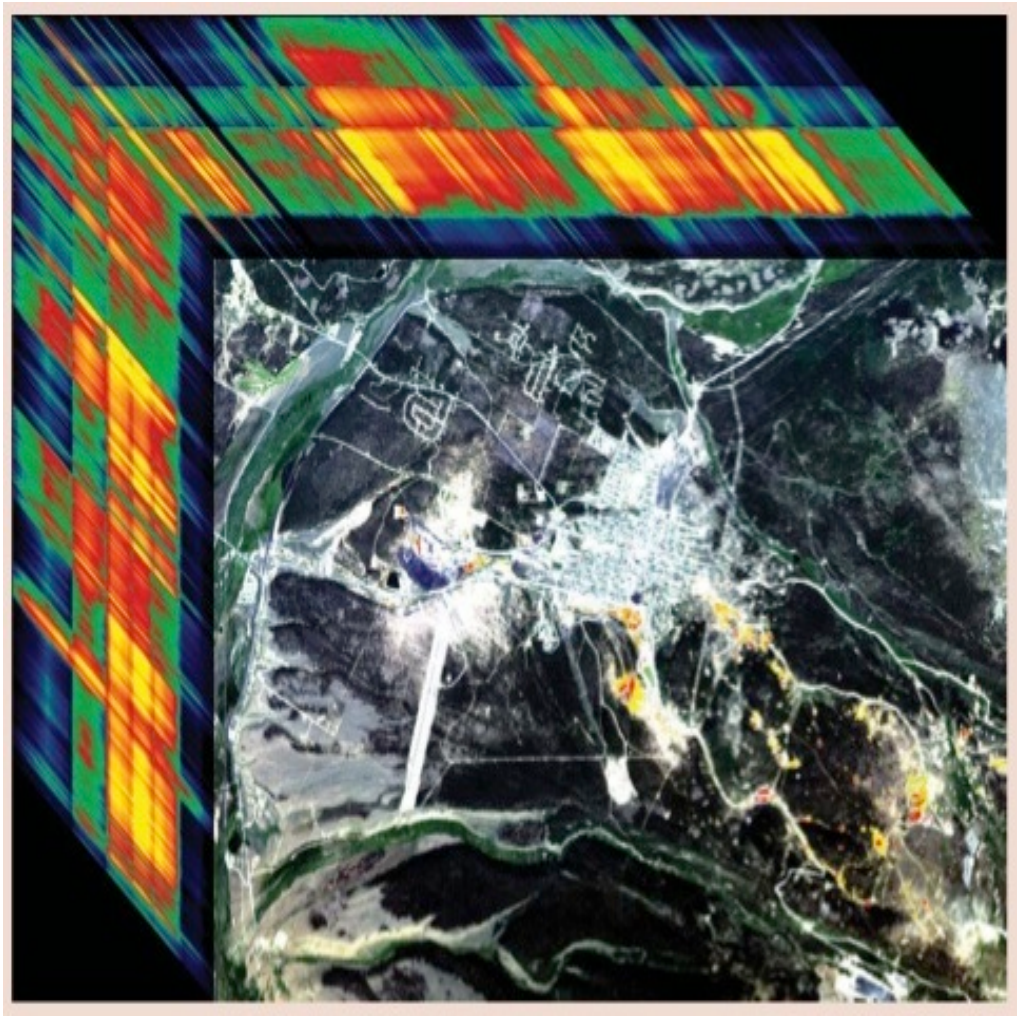
(a) Greyscale-valued image function



(b) Function-valued image function

Example: In multispectral/hyperspectral imaging, u represents the **spectral reflectance density function**. The values $u(x, t_k)$, $t_1 < t_2 < \cdots < t_M$ represent intensities of reflected radiation from point x on ground, as captured by satellite reading, at a discrete set of wavelengths, t_k.

Example 1: Hyperspectral imaging



In practical situations, digital multispectral/hyperspectral images are represented – and quite naturally so – by **vector-valued functions**,

$$u : X \to \mathbb{R}^M,$$

i.e.,

$$u(x) = (u_1(x), u_2(x), \cdots, u_M(x)),$$

where

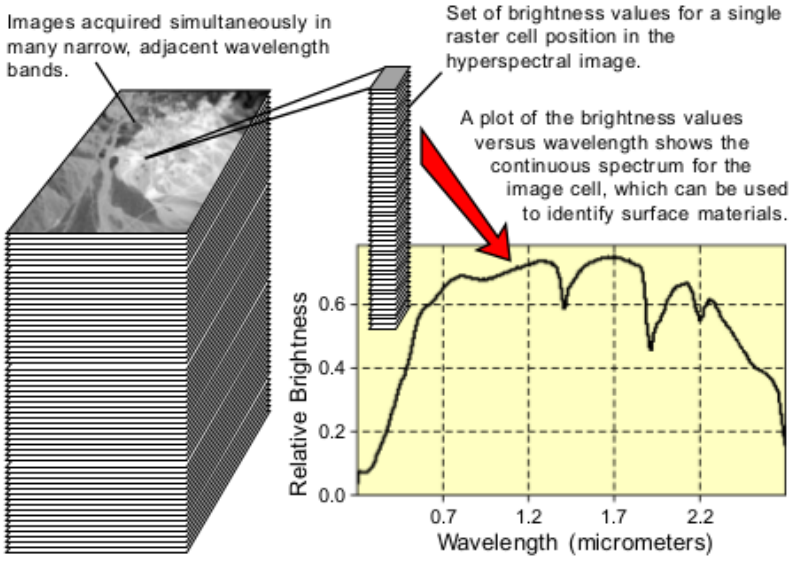
$$u_k : X \to \mathbb{R}, \quad 1 \leq k \leq M$$

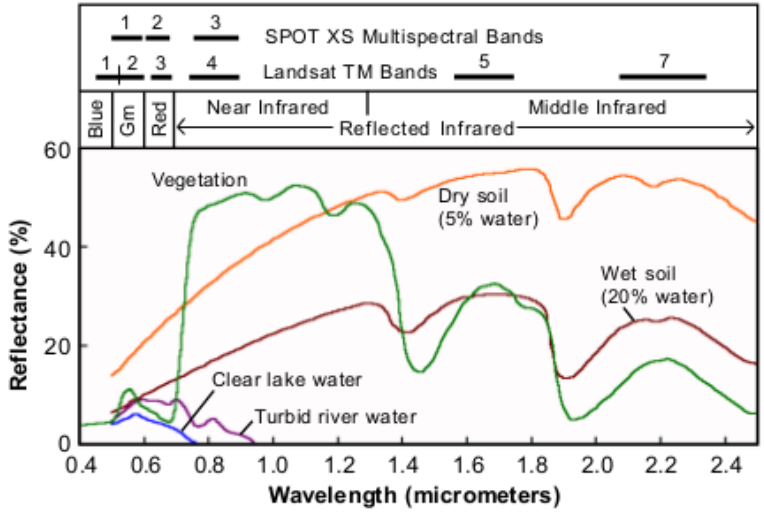
are the usual real-valued **image functions**. (Of course, RGB images are special, low-dimensional, cases.)

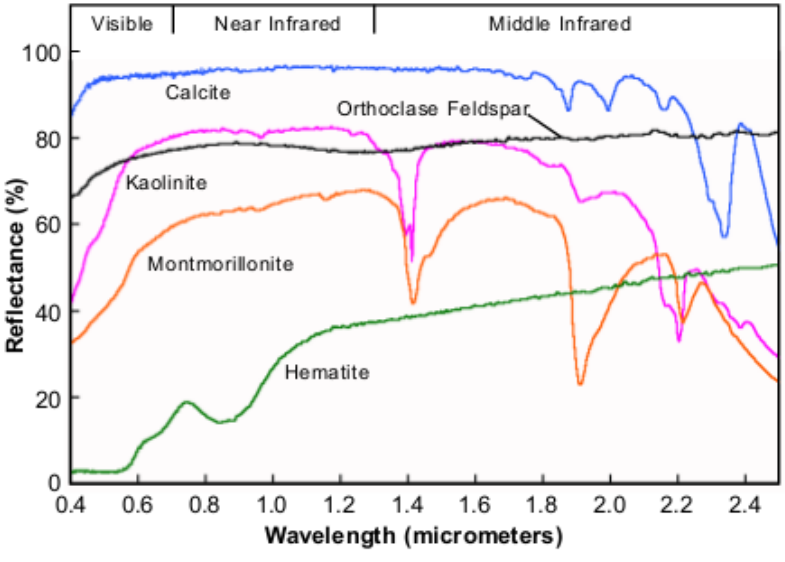
As such, digital hyperspectral images are treated as "data cubes."

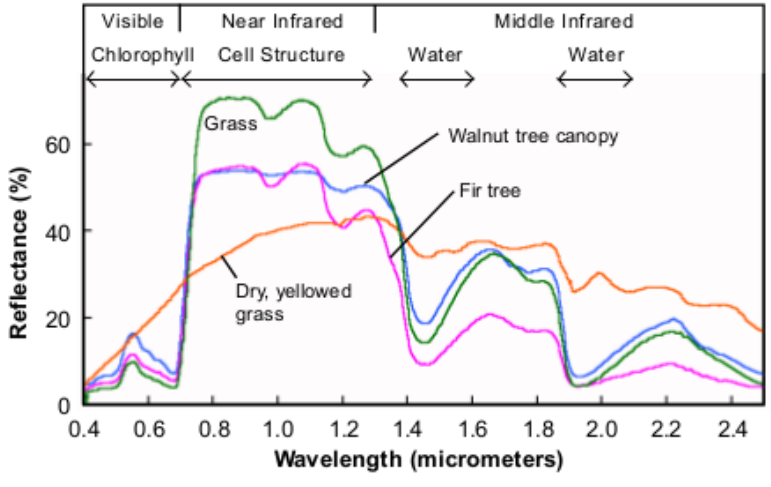
That being said, we believe that it is instructive to start with the continuous, multifunction approach. One advantage is that we may be able to better characterize the nature of spectral functions. No insight is provided by treating spectral functions simply as M-vectors.

The nature of spectral functions

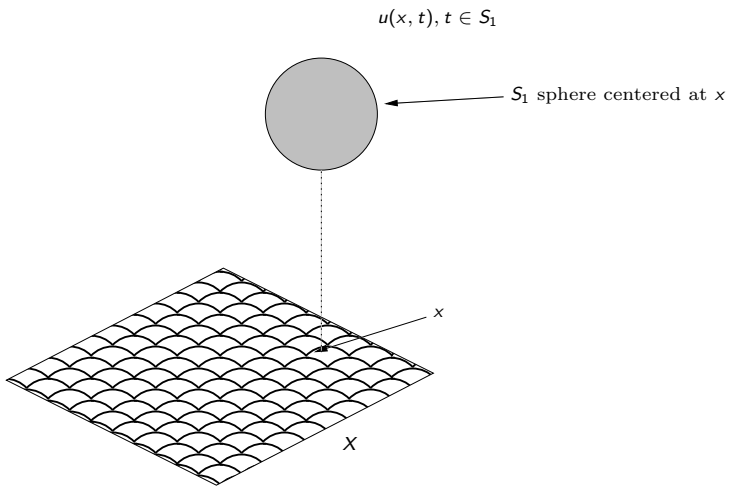








Example 2: High angular resolution diffusion imaging (HARDI)



At each pixel $x \in X$, the HARDI signal $u(x,t)$ is related to probability of diffusion of water molecules in the direction $t \in S^2$ (unit sphere).

u can be represented as a **function-valued image mapping**:

$$u : X \to L^2(\mathbb{S}^2).$$

"q-space imaging"

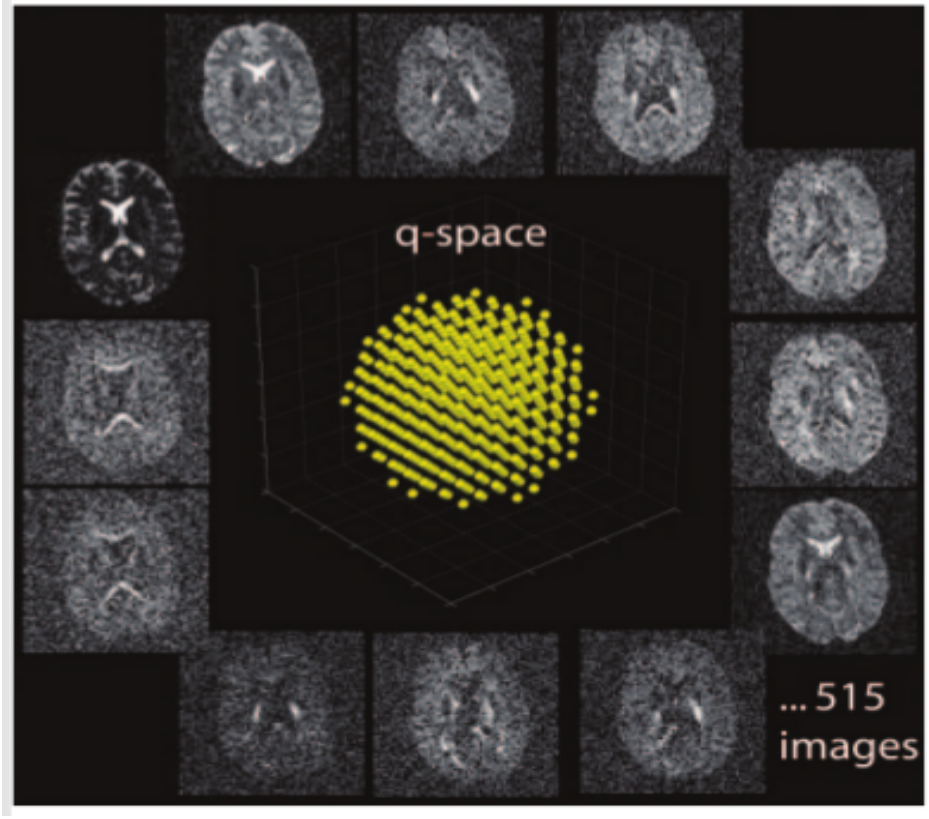


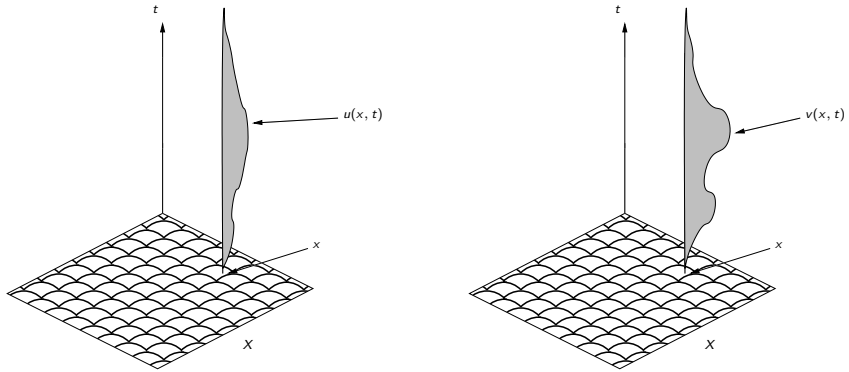
Figure 12. Series of diffusion-weighted MR brain images obtained with variations in the direction and strength of the diffusion gradient in the pulsed gradient SE sequence. Each image shows the signal sampled at one point in q-space (one yellow dot). Every sampling point in q-space corresponds to a specific direction and strength of the diffusion gradient.

From "Understanding Diffusion MR Imaging Techniques: From Scalar Diffusion-weighted imaging to Diffusion Tensor Imaging and Beyond," P. Hagmann, L. Jonasson, P. Maeder, J.-P. Thiran, V.J. Weeden, R. Meuli, Radiographics 2006, 26:S205-S223, Published online 10.1148rg.26si065510.

“Connectome”

In the brain, neuronal tissue is highly fibrillar, consisting of tightly packed and aligned axons that are surrounded by glial cells and organized in bundles. Diffusion of water parallel to fiber is typically greater than perpendicular to it.
In “tractography”, the streamlines associated with the anisotropic diffusion vector fields are constructed in order to give an idea of the connectivity of neurons.

A complete metric space (Y, d_Y) of function-valued images



(a) (b)

Linear space: For $u, v : X \to L^2(R_g)$, define

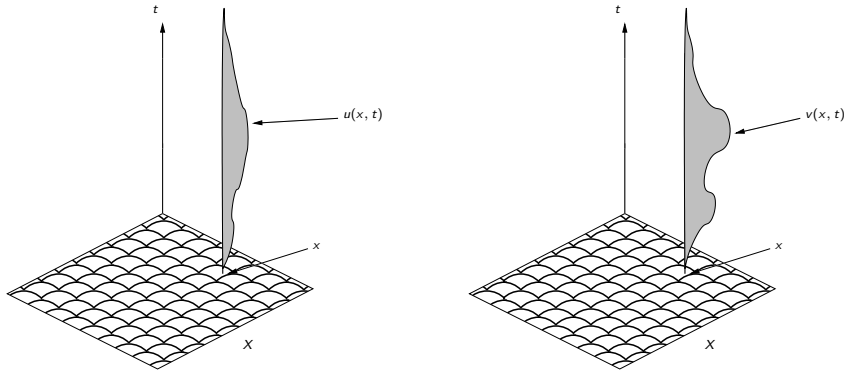
$$(c_1 u + c_2 v)(x,t) = c_1 u(x,t) + c_2 v(x,t), \quad \text{etc. (linear space)}$$

Normed linear space Y: For $u : X \to L^2(R_g)$, norm of $u(x)$ is given by

$$\|u(x)\|^2_{L^2(R_g)} = \int_{R_g} u(x,t)^2 \, dt.$$

Integrate over all $x \in X$ to define norm of u:

$$\|u\|^2_Y = \int_X \|u(x)\|^2_{L^2(R_g)} \, dx.$$



(a) (b)

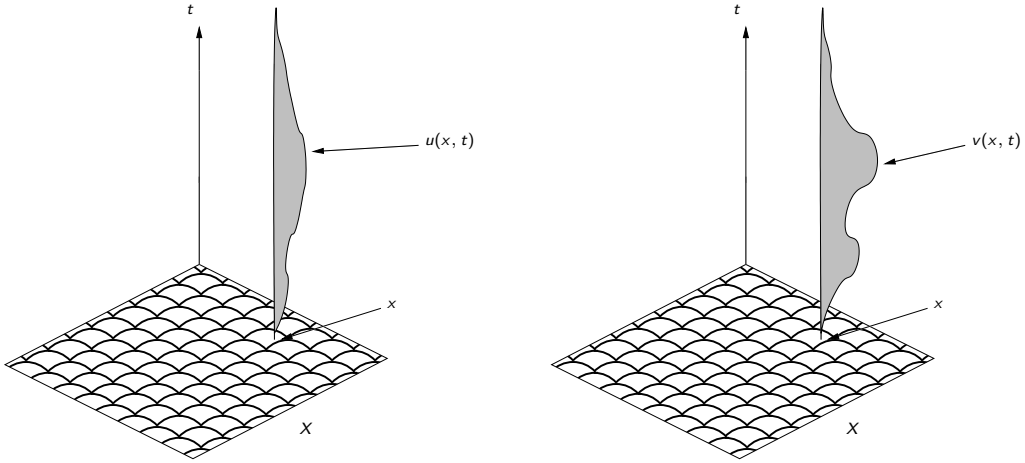
Complete metric space (Y, d_Y):

1. At each $x \in X$, compute L^2 distance between functions $u(x)$ and $v(x)$:

$$\|u(x) - v(x)\|^2_{L^2(R_g)} = \int_{R_g} [u(x,t) - v(x,t)]^2 \, dt$$

2. Integrate over all $x \in X$:

$$[\, d_Y(u,v)\,]^2 = \int_X \|u(x) - v(x)\|^2_{L^2(R_g)} \, dx.$$



(a) (b)

Hilbert space:
Since $u(x), v(x) \in L^2(R_g)$, we may compute their inner product $\langle u(x), v(x) \rangle_{L^2(R_g)}$. Integrate over all $x \in X$ to define inner product between function-valued image mappings,

$$\langle u, v \rangle_Y = \int_X \langle u(x), v(x) \rangle_{L^2(R_g)} \, dx, \quad u, v \in Y.$$

Complete metric space (Y, d_Y) of function-valued image mappings

$$Y = \{u : X \to L^2(R_g) \mid \|u\|_Y < \infty\}$$

where

$$\|u\|_Y^2 = \int_X \|u(x)\|_{L^2(R_g)}^2 \, dx$$

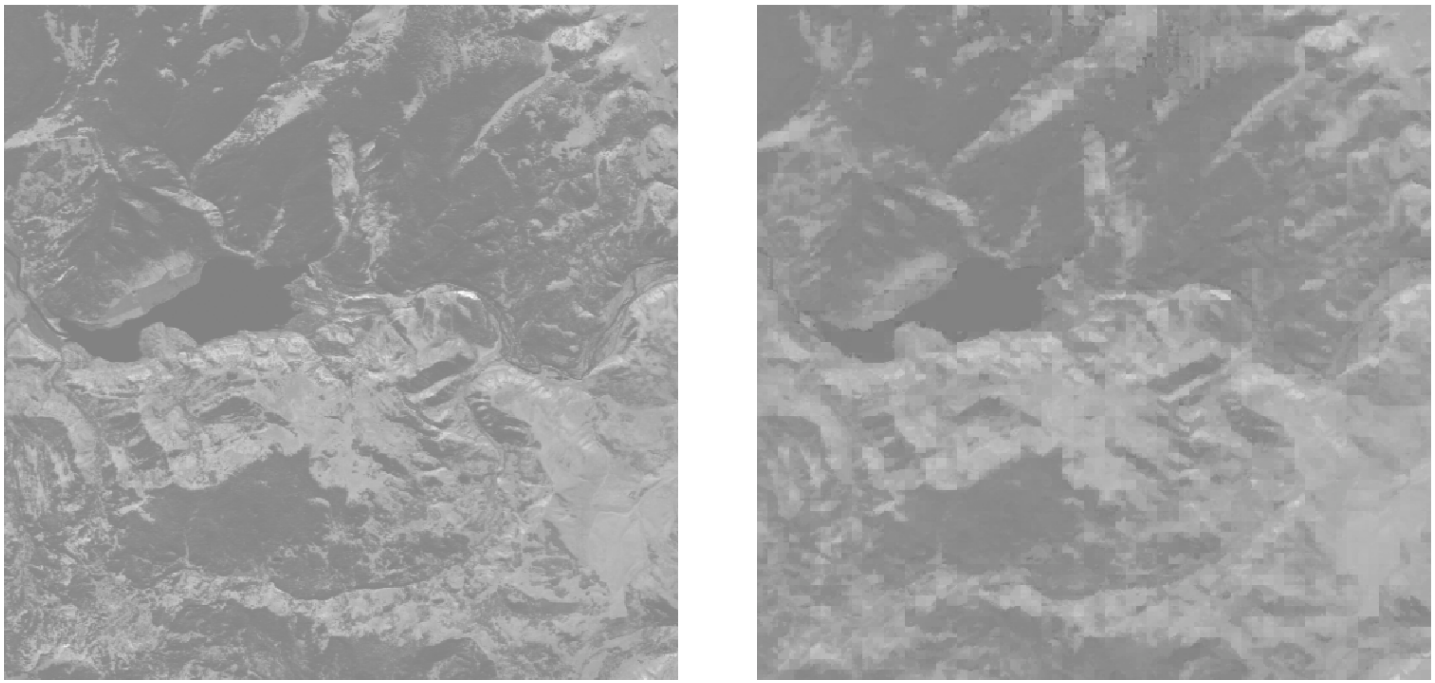
In our applications,

$$R_g = [a, b] \subset R_+ = [0, \infty).$$

We can now go to town and define various fractal transforms on (Y, d_Y) ...

Example: Fractal coding of 224-channel AVIRIS "Yellowstone" image

224-channel AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) hyperspectral image, "Yellowstone calibrated scene 0," a 224-channel image, available from JPL.

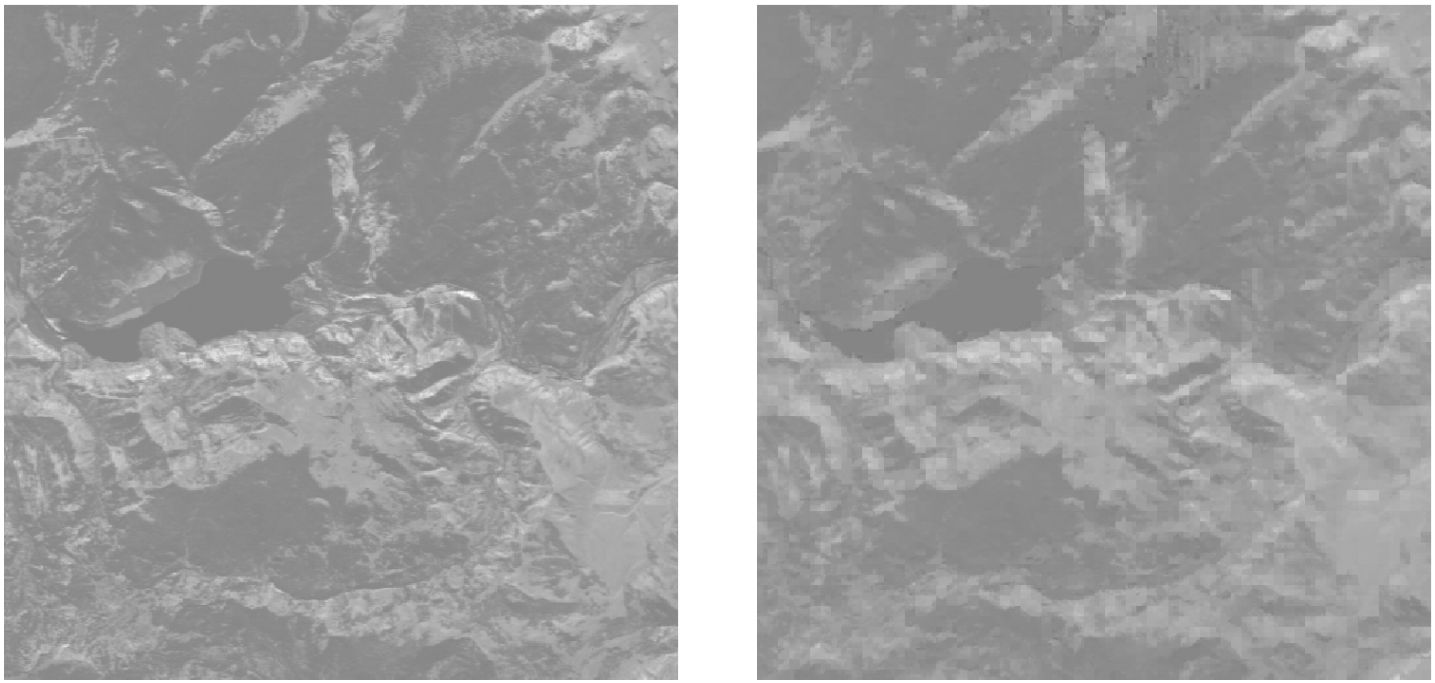


Channel 120. Left: Original. Right: Fractal-based approximation.
8×8-pixel range blocks and 16×16-domain blocks.

(D. La Torre, D. Otero and ERV, Applied Math. Comp. **231**, 435-444 (2014))

Example: Fractal coding of 224-channel AVIRIS "Yellowstone" image

224-channel AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) hyperspectral image, "Yellowstone calibrated scene 0," a 224-channel image, available from JPL.



Channel 220. Left: Original. Right: Fractal-based approximation.
8×8-pixel range blocks and 16×16-domain blocks.

(D. La Torre, D. Otero and ERV, Applied Math. Comp. **231**, 435-444 (2014))

Inverse problems of approximation by fixed points of "nonfractal" contraction mapppings

Recall: Let (Y, d_Y) be a complete metric space and $Con(Y)$ the set of all contraction maps $T : X \to X$. Now let $Con'(Y) \subset Con(Y)$ be a particular class of contraction maps that we wish to consider.

Then given a $y \in Y$ (our "target") and an an $\epsilon > 0$, can we find a $T \in Con'(Y)$ with fixed point $\bar{y}$ such that

$$d_Y(y, \bar{y}) < \epsilon \, ? \tag{47}$$

In other words, can we approximate y with the fixed point $\bar{y}$ to ϵ-accuracy?

Recast this problem as follows:

"Collage coding:" Given a $y \in Y$ and a $\delta > 0$, find T so that

$$d_Y(Ty, y) < \delta \, . \tag{48}$$

So, like, can we find an interesting set of "nonfractal" contraction mappings?

Like, yes!

Warmup:

ODE initial value problem (IVP) for $x(t)$ on $\mathbb{R}$:

$$x' = f(x,t) \quad x(t_0) = x_0\,. \tag{49}$$

For the moment, simply assume that $f(x,t)$ is continuous in x and t.

IVP in (49) is equivalent to the following integral equation,

$$x(t) = x_0 + \int_{t_0}^{t} f(x(s),s)\,ds\,. \tag{50}$$

The solution $x(t)$ may be viewed as the fixed point of the following (Picard) integral operator $T : C(I) \to C^1(I)$, where I is an interval containing t_0:

$$v(t) = (Tu)(t) = x_0 + \int_{t_0}^{t} f(u(s),s)\,ds\,. \tag{51}$$

Continuity of f is sufficient to guarantee the existence of a solution to IVP (49) (Peano's Theorem).

But if f satisfies the following (uniform) Lipschitz property: For all $t \in I$,

$$|f(x_2, t) = f(x_1, t)| \leq K|x_2 - x_1| \quad \forall x_1, x_2 \in J\,, \tag{52}$$

where J is an interval containing x_0, the solution to (49) is **unique** .

Reason: $T : S \to S$ is a contraction mapping, where S is a space of continuous functions.

From Banach's Fixed Point Theorem, there exists a unique $x \in S$ such that

$$x = Tx\,. \tag{53}$$

We have the ingredients for an inverse problem!

Given a function $u(t)$, the evolution of which we suspect to be described by a DE of the form,

$$x' = f(x,t)\,, \tag{54}$$

can we find a function $g(x,t)$ which defines a Picard integral operator T which, in turn, maps x close to itself so that, at least,

$$x' \approx g(x,t)\,. \tag{55}$$

Strategy: We consider a limited class of functions g, e.g. polynomials in x (and perhaps t) and optimize over the coefficients.

Easily extended to systems of ODEs.

First done in:

H. Kunze and ERV, Solving inverse problems for ordinary differential equations using the Picard contraction mapping, Inverse Problems **20**, 3, 977-991 (1999).

Example 1(a): (The second example considered in the 1999 paper.) Let $x(t) = t^2$ be the target solution on $I = [0, 1]$. Find the best ODE of the form

$$\frac{dx}{dt} = c_0 + c_1 x\,, \quad x(0) = x_0\,, \tag{56}$$

with c_0, c_1 and x_0 to be determined.

Result: The IVP

$$\frac{dx}{dt} = \frac{5}{12} + \frac{35}{18}x\,, \quad x(0) = -\frac{1}{27}\,. \tag{57}$$

The solution to this IVP is

$$\bar{x}(t) = \frac{67}{378}\exp\left(\frac{35}{18}t\right) - \frac{3}{14}\,. \tag{58}$$

L^2 distance between target x and $\bar{x}$ is

$$\|x - \bar{x}\|_2 = 0.0123\,. \tag{59}$$

Example 1(b): If we impose the condition that $x(0) = x_0 = 0$, then only c_0 and c_1 are to be determined. The resulting DE is

$$\frac{dx}{dt} = \frac{5}{12} + \frac{35}{16}x\,, \quad x(0) = -\frac{1}{27}\,. \tag{60}$$

The soluton to this IVP is

$$\bar{x}(t) = \frac{1}{7}\exp\left(\frac{35}{16}t\right) - \frac{1}{7}\,. \tag{61}$$

L^2 distance between target x and $\bar{x}$ is

$$\|x - \bar{x}\|_2 = 0.0463\,. \tag{62}$$

Our "collage coding" approach provides a method of performing "parameter estimation" for systems of ODEs

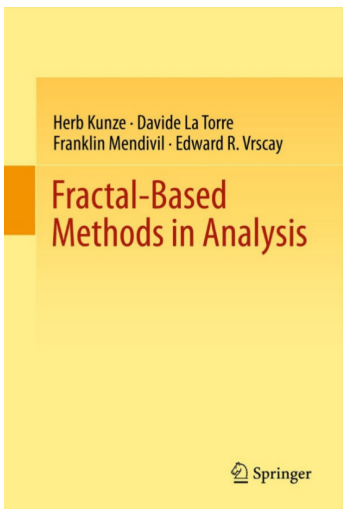
This is important in a number of scientific applications, including

- chemical kinetics, especially "pharmacokinetics,"

Since the 1999 paper, we have built up our "Collage Industry", extending our approach to treat a variety of inverse problems in:

- PDEs (hyperbolic, elliptic, parabolic)
- Random ODEs,
- Stochastic ODEs,
- Random PDEs,
- Stochastic PDEs.

Some of these results (up to 2012) are presented in



Yes! It's over!

May you be rewarded generously for your patience.