Available online at www.sciencedirect.com

SCIENCE DIRECT

Journal of Approximation Theory 141 (2006) 99–117

JOURNAL OF Approximation Theory

www.elsevier.com/locate/jat

Construction of recurrent bivariate fractal interpolation surfaces and computation of their box-counting dimension

P. Bouboulis[a], Leoni Dalla[b], V. Drakopoulos[c,*]

[a] *Department of Informatics and Telecommunications, Telecommunications and Signal Processing, University of Athens, Panepistimioupolis 157 84, Athens, Greece*

[b] *Department of Mathematics, Mathematical Analysis, University of Athens, Panepistimioupolis 157 84, Athens, Greece*

[c] *Department of Informatics and Telecommunications, Theoretical Informatics, University of Athens, Panepistimioupolis 157 84, Athens, Greece*

Received 30 December 2004; accepted 31 January 2006

Communicated by Ulrich Reif

Available online 24 April 2006

Abstract

Recurrent bivariate fractal interpolation surfaces (RBFISs) generalise the notion of affine fractal interpolation surfaces (FISs) in that the iterated system of transformations used to construct such a surface is non-affine. The resulting limit surface is therefore no longer self-affine nor self-similar. Exact values for the box-counting dimension of the RBFISs are obtained. Finally, a methodology to approximate any natural surface using RBFISs is outlined.

Keywords: Fractal interpolation functions; IFS; RIFS; Fractals; Bivariate fractal interpolation surfaces; Box-counting dimension; Minkowski dimension

1. Introduction

Fractal theory has been drawing considerable attention of researchers in various scientific areas. The application of fractals created by *iterated function systems* (IFSs) in the area of image compression is probably the most known one. Applications of fractal surfaces have been also found in computer graphics, metallurgy, geology, chemistry, medical sciences and several other

* Corresponding author.

E-mail addresses: bouboulis@di.uoa.gr (P. Bouboulis), ldalla@math.uoa.gr (Leoni Dalla), vasilios@di.uoa.gr (V. Drakopoulos).

0021-9045/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jat.2006.01.006

areas where there is great need to construct extremely complicated objects; see for example [12,20,18,16].

Mazel and Hayes (see [15,14,19]) used fractal interpolation functions (FIFs) (introduced by Barnsley [1]) to approximate discrete sequences of data (like one-dimensional signals). They demonstrated the effectiveness of their method by modelling seismic and electrocardiogram data. Recently, Navascues and Sebastian (in [17]) gave a generalisation of Hermite functions using FIFs.

Fractal interpolation surfaces (FISs) were used to approximate surfaces of rocks, metals (see [22]), terrains [23], planets [5] and to compress images [4]. The object of this paper is to introduce a new, more general, class of FISs suitable to approximate any natural surface.

Self-affine FISs were first introduced in [13] in the case where the domains are triangular and the interpolation points on the boundary of the domain are coplanar. A few years later Geronimo and Hardin [8] generalised the construction of Massopust to allow for more general boundary data and domains. Simultaneously, in Hardin and Massopust [9], $\mathbb{R}^m$-valued multivariable fractal functions were investigated. The latter two constructions use recurrent iterated function systems (RIFSs).

Some problems in the construction by [13] remained unsolved, amongst which was the lack of free contractivity factors, which are necessary in modelling complicated surfaces. A general constructive method of generating affine FISs is presented in [24]. Xie and Sun [21] and Xie et al. [22] presented a construction of a compact set that contains the interpolation points defined on a rectangular domain. A special construction of bivariate FIS (BFIS) is given in [11]. Dalla [6] gives some conditions in such a way that the bivariate IFS gives a FIS.

In this paper, we introduce recurrent BFISs as a generalisation of the aforementioned methods in order to gain more flexibility in natural-shape generation or in image compression. The main advantage of this new class of FIS is that they are neither self-affine nor self-similar in contrast to all the previously mentioned constructions. Our method is used in [4] for image reconstruction and offers the advantage of a more flexible fractal modelling compared to previous fractal techniques (based on affine transformations). The compression ratio for the aforementioned fractal scheme (though not close enough to JPEG2000) is higher than other fractal methods or JPEG.

2. Recurrent bivariate IFSs on rectangular grids

Any subset of $\mathbb{R}^3$ of the form $[a, b]\times[c, d]\times\mathbb{R}$ may easily be scaled down to $[0, 1]\times[0, p]\times\mathbb{R}$. Therefore, for notation purposes, we may focus on the metric space $X = [0, 1] \times [0, p] \times \mathbb{R}$. Let $\Delta = \{(x_i, y_j, z_{ij}) : i = 0, 1, \ldots, N; j = 0, 1, \ldots, M\}$ be an interpolating set with $(N + 1)\cdot(M + 1)$ *interpolation points*, such that $0 = x_0 < x_1 < \cdots < x_N = 1$ and $0 = y_0 < y_1 < \cdots < y_M = p$. Furthermore let $Q = \{(\hat{x}_k, \hat{y}_l, \hat{z}_{kl}) : k = 0, 1, \ldots, K; l = 0, 1, \ldots, L\}$ be a set with $(K + 1) \cdot (L + 1)$ points with $Q \subset \Delta$ $(Q \neq \Delta)$, such that $0 = \hat{x}_0 < \hat{x}_1 < \cdots < \hat{x}_K = 1$ and $0 = \hat{y}_0 < \hat{y}_1 < \cdots < \hat{y}_L = p$. The interpolation points divide $[0, 1] \times [0, p]$ into $N \cdot M$ rectangles $I_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$, $i = 1, \ldots, N$ and $j = 1, \ldots, M$, which we call *sections* (or) *regions*, while the points of Q divide $[0, 1] \times [0, p]$ to $K \cdot L$ rectangles $J_{kl} = [\hat{x}_{k-1}, \hat{x}_k] \times [\hat{y}_{l-1}, \hat{y}_l]$, $k = 1, \ldots, K$ and $l = 1, \ldots, L$ which we simply call *intervals* (or) *domains*. It is evident that for every interval J_{kl} there are some sections lying inside.

Define a labelling map $\mathbb{J}$: $\{1, 2, \ldots, N\}\times\{1, 2, \ldots, M\} \to \{1, 2, \ldots, K\}\times\{1, 2, \ldots, L\}$ with $\mathbb{J}(i, j) = (k, l)$, such that $\hat{x}_k - \hat{x}_{k-1} > x_i - x_{i-1}$ and $\hat{y}_l - \hat{y}_{l-1} > y_j - y_{j-1}$ for $i = 1, 2, \ldots, N$,

$j = 1, 2, \ldots, M$ and contractive mappings $w_{ij}: X \to X$ satisfying

$$w_{ij}\begin{pmatrix}\hat{x}_{k-1}\\ \hat{y}_{l-1}\\ \hat{z}_{k-1,l-1}\end{pmatrix} = \begin{pmatrix}x_{i-1}\\ y_{j-1}\\ z_{i-1,j-1}\end{pmatrix}, \quad w_{ij}\begin{pmatrix}\hat{x}_{k}\\ \hat{y}_{l-1}\\ \hat{z}_{k,l-1}\end{pmatrix} = \begin{pmatrix}x_{i}\\ y_{j-1}\\ z_{i,j-1}\end{pmatrix},$$
$$w_{ij}\begin{pmatrix}\hat{x}_{k-1}\\ \hat{y}_{l}\\ \hat{z}_{k-1,l}\end{pmatrix} = \begin{pmatrix}x_{i-1}\\ y_{j}\\ z_{i-1,j}\end{pmatrix} \quad \text{and} \quad w_{ij}\begin{pmatrix}\hat{x}_{l}\\ \hat{y}_{l}\\ \hat{z}_{kl}\end{pmatrix} = \begin{pmatrix}x_{i}\\ y_{j}\\ z_{i,j}\end{pmatrix}, \tag{1}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. The w_{ij} map the vertices of the interval $J_{kl} = J_{\mathbb{J}(i,j)}$ to the vertices of the section I_{ij}. Finally, let $\Phi: \{1, \ldots, N\} \times \{1, \ldots, M\} \to \{1, \ldots, N \cdot M\}$ be a 1–1 function (i.e. an enumeration of the set $\{(i, j) : i = 1, \ldots, N; j = 1, \ldots, M\}$).

A *recurrent iterated function system* (RIFS) associated with the set of data Δ consists of the IFS $\{X; w_{i,j}, i = 1, 2, \ldots, N; j = 1, 2, \ldots, M\}$ $\left(\text{or, somewhat more briefly, as } \left\{X; w_{1-N,1-M}\right\}\right)$ together with a row-stochastic matrix $(p_{nm} \in [0, 1] : n, m \in \{\Phi(i, j), i = 1, \ldots, N; j = 1, \ldots, M\})$, such that

$$\sum_{m=1}^{N\cdot M} p_{nm} = 1, \quad n = 1, \ldots, N \cdot M. \tag{2}$$

The recurrent structure is given by the *connection matrix* $C = (C_{nm})$, defined by

$$C_{nm} = \begin{cases} 1 & \text{if } p_{mn} > 0, \\ 0 & \text{if } p_{mn} = 0, \end{cases}$$

for $n, m = 1, 2, \ldots, N \cdot M$. The transition probability for a certain discrete time Markov process is p_{nm}, which gives the probability of transfer into state m given that the process is in state n. Eq. (2) says that whichever state the system is in (say n), a set of probabilities is available that sum to 1, and they describe the possible states to which the system transits at the next step.

In this paper, we study the special case where w_{ij} are transformations of the form

$$w_{ij}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}a_{ij}x + b_{ij}\\ c_{ij}y + d_{ij}\\ e_{ij}x + f_{ij}y + g_{ij}xy + s_{ij}z + k_{ij}\end{pmatrix} = \begin{pmatrix}\phi_{ij}(x)\\ \psi_{ij}(y)\\ F_{ij}(x, y, z)\end{pmatrix}. \tag{3}$$

Define the function T_{ij} by

$$T_{ij}\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}a_{ij}x + b_{ij}\\ c_{ij}y + d_{ij}\end{pmatrix} = \begin{pmatrix}\phi_{ij}(x)\\ \psi_{ij}(y)\end{pmatrix}, \tag{4}$$

so that $w_{ij} = (T_{ij}, F_{ij})^{\mathrm{t}}$ (where A^{t} denotes the transpose matrix of A). From Eq. (3) eight linear equations arise which can always be solved for $a_{ij}, b_{ij}, c_{ij}, d_{ij}, g_{ij}, e_{ij}, f_{ij}, k_{ij}$ in terms of the coordinates of the interpolation points and the *vertical scaling* (or *contractivity*) factor s_{ij} (see [6,21]). We easily find that:

$$a_{ij} = \frac{x_i - x_{i-1}}{\hat{x}_k - \hat{x}_{k-1}},$$

$$b_{ij} = \frac{\hat{x}_k x_{i-1} - \hat{x}_{k-1} x_i}{\hat{x}_k - \hat{x}_{k-1}},$$

$$c_{ij} = \frac{y_j - y_{j-1}}{\hat{y}_l - \hat{y}_{l-1}},$$

$$d_{ij} = \frac{\hat{y}_l y_{j-1} - \hat{y}_{l-1} y_j}{\hat{y}_l - \hat{y}_{l-1}},$$

$$g_{ij} = \frac{z_{ij} + z_{i-1,j-1} - z_{i-1,j} - z_{i,j-1} - s_{ij}\left(\hat{z}_{kl} + \hat{z}_{k-1,l-1} - \hat{z}_{k-1,l} - \hat{z}_{k,l-1}\right)}{\left(\hat{y}_l - \hat{y}_{l-1}\right)\left(\hat{x}_k - \hat{x}_{k-1}\right)},$$

$$e_{ij} = \frac{z_{i-1,j-1} - z_{i,j-1} - s_{ij}\left(\hat{z}_{k-1,l-1} - \hat{z}_{k,l-1}\right) - g_{ij}\hat{y}_{l-1}\left(\hat{x}_{k-1} - \hat{x}_k\right)}{\hat{x}_{k-1} - \hat{x}_k},$$

$$f_{ij} = \frac{z_{i-1,j-1} - z_{i,j-1} - s_{ij}\left(\hat{z}_{k-1,l-1} - \hat{z}_{k-1,l}\right) - g_{ij}\hat{x}_{l-1}\left(\hat{y}_{k-1} - \hat{y}_k\right)}{\hat{y}_{k-1} - \hat{y}_k},$$

$$k_{ij} = z_{ij} - e_{ij}\hat{x}_k - f_{ij}\hat{y}_l - s_{ij}\hat{z}_{kl} - g_{ij}\hat{x}_k\hat{y}_l, \tag{5}$$

If the vertical scaling factors obey $|s_{ij}| < 1$, then there is a metric d on X equivalent to the Euclidean metric, such that w_{ij} is a contraction with respect to d (i.e. $\exists \hat{s}_{ij} : 0 \leqslant \hat{s}_{ij} < 1 : d(w_{ij}(\bar{x}), w_{ij}(\bar{y})) \leqslant \hat{s}_{ij} d(\bar{x}, \bar{y}), \forall \bar{x}, \bar{y} \in X$). One such metric d is given by (see [6]):

$$d\left((x_1, y_1, z_1), (x_2, y_2, z_2)\right) = |x_1 - x_2| + |y_1 - y_2| + \theta\,|z_1 - z_2|,$$

where

$$\theta = \min\left\{\frac{\min_{i,j}\{1 - a_{ij}\}}{\max_{i,j}\{2(|e_{ij}| + p|g_{ij}|)\}}, \frac{\min_{i,j}\{1 - c_{ij}\}}{\max_{i,j}\{2(|f_{ij}| + |g_{ij}|)\}}\right\}.$$

The corresponding RIFS is called *recurrent bivariate IFS* (RBIFS).

It has been proved in [3] that there is a fixed point (attractor) A of this RBIFS. In this case, we can easily see that

$$T_{ij}(J_{\mathbb{J}(i,j)}) = I_{ij}, \quad i = 1, \ldots, N, \quad j = 1, \ldots, M.$$

We also assume that $pN,\ pK \in \mathbb{N}$, the sections (defined by the interpolation points) are squares of side $\delta = 1/N$, while the intervals are squares of side $\psi = 1/K$ (thus $M = pN$, $L = pK$) and the number

$$a = \frac{\psi}{\delta} = \frac{N}{K}$$

is an integer greater than one. The number a^2 expresses how many sections lie inside any interval.

If we define the enumeration $\Phi(i, j) = (i - 1)M + j,\ i = 1, \ldots, N$ and $j = 1, \ldots, M$, then $\Phi^{-1}(n) = ((n-1) \operatorname{div} M + 1,\ (n-1) \mod M + 1),\ n = 1, \ldots, N \cdot M$ and the $NM \times NM$ stochastic matrix (p_{nm}) is defined by

$$p_{nm} = \begin{cases} \frac{1}{q_n} & \text{if } I_{\Phi^{-1}(n)} \subseteq J_{\mathbb{J}(\Phi^{-1}(m))}, \\ 0 & \text{otherwise,} \end{cases}$$

where q_n is the number of non-zero elements of the nth row of the stochastic matrix (p_{nm}). This means that p_{nm} is positive iff there is a transformation T_{ij}, which maps an interval containing the nth section (i.e. $I_{\Phi^{-1}(n)} = I_{(n-1)\operatorname{div} M+1,(n-1) \mod M+1}$) to the mth section

(i.e. $I_{(m-1)\,\mathrm{div}\,M+1,(m-1)\bmod M+1}$). Let us take a point in $I_{ij}\times\mathbb{R}$, $i=1,\ldots,N$, $j=1,\ldots,M$. We say that we are in state n, if $n=\Phi(i,j)$. The matrix (p_{nm}) shows the probability of applying the map $w_{\Phi^{-1}(m)}$ to that point, so that the system transits to state m. Finally, we define the enumeration $\hat{\Phi}(k,l)=(k-1)K+l$, $k=1,\ldots,K, l=1,\ldots,L$ and the *connection vector* $C^{\mathrm{v}}=\{c_1^{\mathrm{v}},c_2^{\mathrm{v}},\ldots,c_{NM}^{\mathrm{v}}\}$ as follows:

$$c_n^{\mathrm{v}}=\hat{\Phi}(\mathbb{J}(\Phi^{-1}(n))),\quad n=1,2,\ldots,NM.$$

If the attractor A of the above RBIFS is the graph of a continuous function, then it is called a *recurrent bivariate fractal interpolation surface* (RBFIS). In the next section, we compute its box-counting dimension.

3. The box-counting dimension of RBFIS

Let B be any non-empty compact subset of $\mathbb{R}^3$ and let $\mathcal{N}(B,\varepsilon)$ be the *smallest number of* (*closed*) *balls of radius* ε that cover B. Let

$$\underline{D}=\underline{D}(B)=\liminf_{\varepsilon>0}\frac{\log\mathcal{N}(B,\varepsilon)}{\log(1/\varepsilon)}\quad\text{and}\quad\overline{D}=\overline{D}(B)=\limsup_{\varepsilon>0}\frac{\log\mathcal{N}(B,\varepsilon)}{\log(1/\varepsilon)}$$

be the *lower* and *upper box-counting dimension* of B, respectively; if

$$D=D(B)=\lim_{\varepsilon\to0}\frac{\log\mathcal{N}(B,\varepsilon)}{\log(1/\varepsilon)}$$

exists, then D is called the *box-counting dimension* of B (see [2]).

In the latter case, we will use the notation $D=D(B)=\underline{D}(B)=\overline{D}(B)$ and will say "B has box-counting dimension D". In practice, we usually use covers of closed cubes of side length $(1/k^n)$, where $k\in\mathbb{N}$. If $\mathcal{N}_n(B)$ denotes the *number of cubes of side* $1/k^n$ that intersect B and

$$D=\lim_{n\to\infty}\frac{\log\mathcal{N}_n(B)}{\log k^n}$$

exists then B has box-counting dimension D. To compute $D(B)$ we usually use covers that differ from those above. Assume that one uses covers from the set $\{\mathcal{C}_\varepsilon:\varepsilon>0\}$, which is formed by *sets of radius* ε and let $\mathcal{N}^*(\varepsilon)$ be the minimum number of sets in $\mathcal{C}_\varepsilon$ that cover B. If we can find constant numbers c_1 and c_2, such that $c_1\mathcal{N}(B,\varepsilon)\leqslant\mathcal{N}^*(\varepsilon)\leqslant c_2\mathcal{N}(B,\varepsilon)$, then $\mathcal{N}^*(\varepsilon)$ can replace $\mathcal{N}(B,\varepsilon)$ in the calculation of $D(B)$ (the proof is straightforward).

We will compute the box-counting dimension of the attractor A of the bivariate RIFS on $\mathbb{R}^3$ defined above, in the case that A is the graph of a continuous function. The following definition gives the cover we use to compute the box-counting dimension of this attractor.

Definition 1. Let $r>0$ and

$$C=\left\{\left[\frac{k-1}{a^r},\frac{k}{a^r}\right]\times\left[\frac{l-1}{a^r},\frac{l}{a^r}\right]\times\left[b,b+\frac{1}{a^r}\right]\ :\ k,l,r\in\mathbb{N},\ b\in\mathbb{R}\right\},$$

where C contains overlapping $1/a^r$-mesh cubes. We define,

$\mathcal{N}^*(r)$ as the minimum number of cubes in C necessary to cover A and

$\mathcal{N}(r)$ as the smallest number of $\dfrac{1}{a^r}$ -mesh cubes which cover A.

We may easily deduce the following.

Lemma 1. $\mathcal{N}(r) \leqslant \mathcal{N}^*(r) \leqslant 4\mathcal{N}(r)$.

Definition 2. The maximum range of a function h inside the rectangle $I \subset \mathbb{R}^2$ is defined by

$$\mathcal{R}_h[I] = \max\{|h(x_1, y_1) - h(x_2, y_2)| : (x_1, y_1),\ (x_2, y_2) \in I\}.$$

Lemma 2. *Let the RBIFS be as defined above and suppose that its attractor A is the graph of a continuous function h defined on* $I = [\alpha, \beta] \times [\gamma, \delta] \subseteq \mathbb{R}^2$, $\alpha, \gamma \geqslant 0$. *Also, let w be one of the maps that define the RBIFS with the form*

$$w\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax + b \\ cy + d \\ ex + fy + gxy + sz + k \end{pmatrix} = \begin{pmatrix} \phi(x) \\ \psi(y) \\ F(x, y, z) \end{pmatrix} \tag{6}$$

with

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + b \\ cy + d \end{pmatrix} = \begin{pmatrix} \phi(x) \\ \psi(y) \end{pmatrix}. \tag{7}$$

Then

$$\mathcal{R}_h[T(I)] \leqslant |e|(\beta - \alpha) + |f|(\delta - \gamma) + |g|\delta(\beta - \alpha) + |g|\beta(\delta - \gamma) + |s|\mathcal{R}_h[I].$$

Proof. Let $(x_1, y_1, z_1)^{\mathrm{t}}, (x_2, y_2, z_2)^{\mathrm{t}} \in A$ and

$$(x_1', y_1', z_1')^{\mathrm{t}} = w((x_1, y_1, z_1)^{\mathrm{t}}),$$

$$(x_2', y_2', z_2')^{\mathrm{t}} = w((x_2, y_2, z_2)^{\mathrm{t}}).$$

Then we have

$$\begin{aligned} |z_2' - z_1'| &\leqslant |e||x_2 - x_1| + |f||y_2 - y_1| + |g||x_2y_2 - x_1y_1| + |s||z_2 - z_1| \\ &= |e||x_2 - x_1| + |f||y_2 - y_1| + |g||x_2y_2 - x_2y_1 + x_2y_1 - x_1y_1| + |s||z_2 - z_1| \\ &\leqslant |e||x_2 - x_1| + |f||y_2 - y_1| + |g||y_1||x_2 - x_1| + |g||x_2||y_2 - y_1| + |s||z_2 - z_1| \\ &\leqslant |e|(\beta - \alpha) + |f|(\delta - \gamma) + |g|\delta(\beta - \alpha) + |g|\beta(\delta - \gamma) + |s||z_2 - z_1|. \end{aligned}$$

So, we have the result. □

Define $[\cdot]: \mathbb{R} \to \mathbb{Z}$ to be the greatest integer function.

Lemma 3. *Let* $x, x_1, x_2, \ldots, x_k \in \mathbb{R}$ *be such that* $x_1 + x_2 + \cdots + x_k = \lambda x$. *Then the following inequalities hold*:

$$\lambda x - k \leqslant [x_1] + [x_2] + \cdots + [x_k] \leqslant \lambda x.$$

Definition 3. Let $U, V \in \mathbb{R}^N$ with $U = (u_1, u_2, \ldots, u_N)$ and $V = (v_1, v_2, \ldots, v_N)$. Define the relation $\prec$ as follows:

$$U \prec V \quad \text{iff } u_i < v_i, \quad i = 1, 2, \ldots, N.$$

Definition 4. Let $P = \{(x_i, y_j, z_{ij});\ i = 1, 2 \ldots, N;\ j = 1, 2, \ldots, M\}$ be a set of points in $\mathbb{R}^3$. We call the points of P *x-collinear* iff all the points with the same x coordinate are collinear. We call the points of P *y-collinear* iff all the points with the same y coordinate are collinear.

Theorem 1. *Let the above RBIFS be defined by an irreducible connection matrix C. Let S be the $N \cdot M \times N \cdot M$ diagonal matrix*

$$S = \operatorname{diag}(|s_{\Phi^{-1}(1)}|, |s_{\Phi^{-1}(2)}|, \ldots, |s_{\Phi^{-1}(NM)}|),$$

with $0 < |s_{ij}| < 1,\ i = 1, \ldots, N$ and $j = 1, \ldots, M$. Suppose that the attractor A of the RBIFS is the graph of a continuous function f that interpolates Δ and that the interpolation points of every interval are not x-collinear or not y-collinear. Then, the box-counting dimension of A is given by

$$D(A) = \begin{cases} 1 + \log_a \lambda & \text{if } \lambda > a, \\ 2 & \text{if } \lambda \leqslant a, \end{cases}$$

where $\lambda = \rho(SC) > 0$, the spectral radius of the irreducible matrix $S \cdot C$.

Proof. Let $R_{kl} = R_f[J_{kl}]$ denote the maximum range of f inside the interval J_{kl}. The connection matrix $C = (C_{nm})$ of the RBIFS we mentioned earlier is defined by $C_{nm} = 1$, if $p_{mn} > 0$, 0 otherwise. If $U = (u_1, u_2, \ldots, u_{NM})^{\mathrm{t}} \in \mathbb{R}^{NM}$, we define

$$\Omega(U) = u_1 + u_2 + \cdots + u_{NM}.$$

According to the hypothesis, we may assume that, for every interval J_{kl}, there is a number i_0 such that the points $\{(x_{i_0}, y_{j_v}, z_{i_0,j_v}) : (x_{i_0}, y_{j_v}) \in J_{kl}, v = 1, 2, \ldots, a+1\}$ are not collinear or there is a number j_0 such that the points $\{(x_{i_v}, y_{j_0}, z_{i_v,j_0}) : (x_{i_v}, y_{j_0}) \in J_{kl}, v = 1, 2, \ldots, a+1\}$ are not collinear. Assume all points are not collinear with each other in that sense and let V_{kl} denote their maximum vertical distance from the line defined by the endpoints $(x_{i_0}, y_{j_1}, z_{i_0,j_1})$ and $(x_{i_0}, y_{j_{a+1}}, z_{i_0,j_{a+1}})$ in the first case or by $(x_{i_1}, y_{j_0}, z_{i_1,j_0})$ and $(x_{i_{a+1}}, y_{j_0}, z_{i_{a+1},j_0})$ for the second case. By the term *vertical distance* we mean a distance that is computed only with respect to the z-axis. We will refer to the vertical distance as "height".

After the first iteration (after applying each w_{ij} to the interpolation points lying inside interval $J_{\mathbb{J}(i,j)}$) we obtain $(a+1)^2$ new points and one height $|s_{ij}|V_{\mathbb{J}(i,j)}$ inside every section. To compute the height we use the property of bivariate functions that each vertical line (parallel to zz') is mapped to a vertical line scaled by the vertical scaling factor, along with the property that a bivariate function maps any line parallel to the xz plane (or to the yz plane) to a line also parallel to the xz plane or to yz plane, respectively. Recall that $I = [0, 1] \times [0, p]$. From Lemma 2, we find that the maximum range inside each section will be

$$R_f[I_{ij}] \leqslant \beta_{ij} \frac{a}{N} + |s_{ij}| R_f[J_{\mathbb{J}(i,j)}],$$

where $\beta_{ij} = |e_{ij}| + |f_{ij}| + |g_{ij}| + |g_{ij}|p,\ i = 1, \ldots, N,\ j = 1, \ldots, M$.

Define non-negative vectors B, H_1, R, U_1 and $\mathbb{I}$, by

$$B = \begin{pmatrix} \beta_{\Phi^{-1}(1)} \\ \beta_{\Phi^{-1}(2)} \\ \vdots \\ \beta_{\Phi^{-1}(NM)} \end{pmatrix}, \quad H_1 = \begin{pmatrix} |s_{\Phi^{-1}(1)}| V_{\mathbb{J}\circ\Phi^{-1}(1)} \\ |s_{\Phi^{-1}(2)}| V_{\mathbb{J}\circ\Phi^{-1}(2)} \\ \vdots \\ |s_{\Phi^{-1}(NM)}| V_{\mathbb{J}\circ\Phi^{-1}(NM)} \end{pmatrix},$$

$$R = \begin{pmatrix} |s_{\Phi^{-1}(1)}| R_{\mathbb{J}\circ\Phi^{-1}(1)} \\ |s_{\Phi^{-1}(2)}| R_{\mathbb{J}\circ\Phi^{-1}(2)} \\ \vdots \\ |s_{\Phi^{-1}(NM)}| R_{\mathbb{J}\circ\Phi^{-1}(NM)} \end{pmatrix},$$

$$U_1 = \frac{a}{N} B + R \quad \text{and} \quad \mathbb{I} = (1, 1, \ldots, 1)^{\mathrm{t}}.$$

Since A is the graph of a continuous function defined on $[0, 1] \times [0, p]$, in order to cover the part of A lying inside $I_{ij} \times \mathbb{R}$ we need more cubes than those needed to cover the height $|s_{ij}| V_{\mathbb{J}(i,j)}$ and less cubes than those needed to cover the parallelipiped $I_{ij} \times [z_{\max}[i, j],\ z_{\min}[i, j]]$, where $z_{\max}[i, j],\ z_{\min}[i, j]$ denote the maximum and minimum values of f inside section I_{ij}. Therefore,

$$\sum_{i=1}^{N} \sum_{j=1}^{M} \left[|s_{ij}| V_{\mathbb{J}(i,j)} a^r \right] \leqslant \mathcal{N}^*(r)$$
$$\leqslant \sum_{i=1}^{N} \sum_{j=1}^{M} \left(\left[\left(\beta_{ij} \frac{a}{N} + |s_{ij}| R_{\mathbb{J}(i,j)} \right) a^r \right] + 1 \right) \left(\left[\frac{1}{N} a^r \right] + 1 \right)^2.$$

Hence (from Lemma 3) we have:

$$\sum_{i=1}^{N} \sum_{j=1}^{M} \left(|s_{ij}| V_{\mathbb{J}(i,j)} a^r \right) - NM \leqslant \mathcal{N}^*(r)$$
$$\leqslant \sum_{i=1}^{N} \sum_{j=1}^{M} \left(\left(\beta_{ij} \frac{a}{N} + |s_{ij}| R_{\mathbb{J}(i,j)} \right) a^r + 1 \right) \left(\left[\frac{1}{N} a^r \right] + 1 \right)^2$$

or

$$\Omega(H_1 a^r) - NM \leqslant \mathcal{N}^*(r) \leqslant \Omega\left(U_1 \cdot a^r + \mathbb{I}\right) \cdot \left(\left[\frac{1}{N} a^r \right] + 1 \right)^2.$$

After the second iteration we obtain a^2 squares of side $\frac{1}{aN}$ which we call *sectors*. Each sector is produced from each section lying inside the interval that is mapped to the original section. The maximum ranges produced inside each sector are contained (as coordinates) in the vector

$$U_2 = SC \cdot U_1 + a^2 B \frac{1}{N}$$

while the heights produced inside each sector (a^2 for each section) are contained (as coordinates) in the vector

$$H_2 = SC \cdot H_1.$$

Thus,

$$\Omega\left(H_2 \cdot a^r\right) - a^2 MN \leqslant \mathcal{N}^*(r) \leqslant \Omega\left(U_2 \cdot a^r + a^2 \mathbb{I}\right) \cdot \left(\left[\frac{1}{aN} a^r\right] + 1\right)^2.$$

Therefore, at the κth iteration we will have $a^{2(\kappa-1)}$ ranges and sectors of side $\frac{1}{a^{\kappa-1}N}$ inside each section and using Lemma 3 we obtain:

$$\Omega\left(H_\kappa \cdot a^r\right) - MN a^{2(\kappa-1)} \leqslant \mathcal{N}^*(r) \leqslant \Omega\left(U_\kappa \cdot a^r + \mathbb{I} \cdot a^{2(\kappa-1)}\right) \cdot \left(\left[\frac{1}{a^{\kappa-1}N} a^r\right] + 1\right)^2, \tag{8}$$

where

$$U_\kappa = SC \cdot U_{\kappa-1} + a^{2(\kappa-1)} B \frac{1}{a^{\kappa-2}N} = SC \cdot U_{\kappa-1} + B\frac{a^\kappa}{N}$$

and

$$H_\kappa = SC \cdot H_{\kappa-1}.$$

We can easily prove that

$$U_\kappa = (SC)^{\kappa-1} \cdot R + (SC)^{\kappa-1} B \frac{a}{N} + (SC)^{\kappa-2} \cdot B \frac{a^2}{N} + (SC)^{\kappa-3} \cdot B \frac{a^3}{N} + \cdots + SC \cdot B \frac{a^{\kappa-1}}{N} + B\frac{a^\kappa}{N}$$

and

$$H_\kappa = (SC)^{\kappa-1} \cdot H_1.$$

This holds only if the number κ of the steps is such that $\frac{1}{a^{\kappa-1}N} \geqslant \frac{1}{a^r}$, so that the cubes we use do not intersect. We choose $\kappa \in \mathbb{N}$ such that $r - \mu - 1 \leqslant \kappa < r - \mu$, where $\mu = \frac{\log N}{\log a} - 1 > 0$.

Since $S \cdot C$ is a *non-negative irreducible* matrix, Frobenius' Theorem (see [7, pp. 50–66,10, pp. 542–551]) implies that there is a "unique" strictly positive eigenvector of SC which corresponds to an eigenvalue $\lambda = \rho(SC) > 0$. (Unique in the sense that an irreducible non-negative matrix cannot have two linearly independent non-negative eigenvectors. Thus, any other strictly positive eigenvector has to be a multiple of the first).

We choose the (corresponding to λ) eigenvectors

$$\hat{U} = (\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_{NM}) \quad \text{such that } 0 \prec \hat{U} \prec H_1, \tag{9}$$

$$U^* = (u_1^*, u_2^*, \ldots, u_{NM}^*) \quad \text{such that } U^* \succ U_1 \text{ and } U^* \succ \frac{1}{N} R. \tag{10}$$

Thus, from (8), we have

$$\begin{aligned}\mathcal{N}^*(r) &\leqslant \Omega\left(U_\kappa \cdot a^r + a^{2(\kappa-1)} \mathbb{I}\right) \cdot \left(\left[\frac{1}{a^{\kappa-1}N} a^r\right] + 1\right)^2 \\ &\leqslant \Omega\left(U_\kappa \cdot a^r + a^{2\kappa-2} \mathbb{I}\right) \cdot (a+1)^2 \quad \left(\text{as } a^\mu = \frac{N}{a} \text{ and } \kappa \geqslant r - \mu - 1\right),\end{aligned}$$

$$\begin{aligned}\mathcal{N}^*(r) &\leqslant \Omega\Big((SC)^{\kappa-1}Ra^r+(SC)^{\kappa-1}B\frac{a^{r+1}}{N}+(SC)^{\kappa-2}B\frac{a^{r+2}}{N}+\cdots+(SC)B\frac{a^{r+\kappa-1}}{N}\\ &\quad+B\frac{a^{r+\kappa}}{N}+a^{2\kappa-2}\mathbb{I}\Big)\cdot(a+1)^2 \qquad (11)\\ &\leqslant \Omega\Big((SC)^{\kappa-1}U^*a^r+(SC)^{\kappa-1}U^*a^{r+1}+(SC)^{\kappa-2}U^*a^{r+2}\\ &\quad+\cdots+(SC)U^*a^{r+\kappa-1}+U^*a^{r+\kappa}+a^{2\kappa-2}\mathbb{I}\Big)\cdot(a+1)^2\\ &\quad\Big(\text{as}\frac{1}{N}\,B\prec U^*\text{ and }R\prec U^*\Big)\\ &=\Big(\lambda^{\kappa-1}\Omega(U^*)a^r+\lambda^{\kappa-1}\Omega(U^*)a^{r+1}+\lambda^{\kappa-2}\Omega(U^*)a^{r+2}\\ &\quad+\cdots+\lambda\Omega(U^*)a^{r+\kappa-1}+\Omega(U^*)a^{r+\kappa}+a^{2\kappa-2}\Omega(\mathbb{I})\Big)\cdot(a+1)^2\\ &\quad\text{(as }U^*\text{is an eigenvector).}\end{aligned}$$

We set $\gamma^*=\Omega(U^*)$; thus since $\kappa\leqslant r-\mu$, we have

$$\begin{aligned}\mathcal{N}^*(r) &\leqslant \Big(\lambda^{r-\mu-1}\gamma^*a^r+\lambda^{r-\mu-1}\gamma^*a^{r+1}+\lambda^{r-\mu-2}\gamma^*a^{r+2}\\ &\quad+\cdots+\lambda\gamma^*a^{2r-\mu-1}+\gamma^*a^{2r-\mu}+a^{2r-2\mu-2}NM\Big)\cdot(a+1)^2\,. \qquad (12)\end{aligned}$$

Assuming $\lambda>a$, we have that

$$\mathcal{N}^*(r)\leqslant\lambda^{r-\mu-1}a^r\gamma^*\left(1+\frac{a^{r-2\mu-2}NM}{\lambda^{r-\mu-1}}+\frac{a\left(1-\left(\frac{a}{\lambda}\right)^{r-\mu}\right)}{1-\frac{a}{\lambda}}\right)\cdot(a+1)^2\,,$$

where

$$\left(1+\frac{a^{r-2\mu-2}NM}{\lambda^{r-\mu-1}}+\frac{a\left(1-\left(\frac{a}{\lambda}\right)^{r-\mu}\right)}{1-\frac{a}{\lambda}}\right)>0.$$

For $\lambda>a$ we deduce that:

$$\begin{aligned}\frac{\log\mathcal{N}^*(r)}{r\log a} &\leqslant \frac{(r-\mu-1)\log\lambda}{r\log a}+\frac{r\log a}{r\log a}+\frac{\log\gamma^*}{r\log a}\\ &\quad+\frac{\log\left(1+\frac{a^{r-2\mu-2}NM}{\lambda^{r-\mu-1}}+\frac{a\left(1-\left(\frac{a}{\lambda}\right)^{r-\mu}\right)}{1-\frac{a}{\lambda}}\right)}{r\log a}+\frac{2\log(a+1)}{r\log a}.\end{aligned}$$

Therefore,

$$\overline{D}(G)=\limsup_{r\to\infty}\frac{\log\mathcal{N}^*(r)}{r\log a}\leqslant 1+\log_a\lambda. \qquad (13)$$

If $\lambda\leqslant a$, then from (12) we have

$$\begin{aligned}\mathcal{N}^*(r) &\leqslant (\lambda^{r-\mu-1}\gamma^*a^r+\lambda^{r-\mu-1}\gamma^*a^{r+1}+\lambda^{r-\mu-2}\gamma^*a^{r+2}\\ &\quad+\cdots+\lambda\gamma^*a^{2r-\mu-1}+\gamma^*a^{2r-\mu}+a^{2r-2\mu-2}NM)\cdot(a+1)^2\\ &\leqslant a^{2r-\mu-2}\gamma^*\Big(a+(r-\mu)a^2+NM\Big)(a+1)^2\,.\end{aligned}$$

Hence

$$\frac{\log \mathcal{N}^*(r)}{r \log a} \leqslant \frac{(2r-\mu-2)\log a}{r \log a} + \frac{\log\left(\gamma^*\left(a+(r-\mu)a^2+NM\right)(a+1)^2\right)}{r \log a}$$

and $D(A) \leqslant 2$. But, since A is a continuous surface, we conclude that $D(A) = 2$.

Proceeding similarly, from (8) we have, for $\lambda > a$,

$$\mathcal{N}^*(r) \geqslant \Omega\left(H_\kappa \cdot a^r\right) - MNa^{2\kappa-2},$$

$$\mathcal{N}^*(r) \geqslant \Omega\left((SC)^{\kappa-1} \cdot H_1 \cdot a^r\right) - MNa^{2\kappa-2}. \tag{14}$$

Hence,

$$\begin{aligned}
\mathcal{N}^*(r) &\geqslant \Omega\left((SC)^{\kappa-1} \cdot \hat{U} \cdot a^r\right) - MNa^{2\kappa-2} \quad (\text{since } H_1 \succ \hat{U}) \\
&\geqslant \Omega\left(\lambda^{\kappa-1} \cdot \hat{U} \cdot a^r\right) - MNa^{2\kappa-2} \quad (\hat{U} \text{ is an eigenvector}) \\
&= \lambda^{\kappa-1} a^r \sum_{n=1}^{NM} \hat{u_i} - MNa^{2\kappa-2} \\
&= \lambda^{\kappa-1} a^r \hat{\gamma} - MNa^{2\kappa-2} \quad \left(\hat{\gamma} = \sum_{n=1}^{NM} \hat{u_i}\right) \\
&\geqslant \lambda^{r-\mu-2} \hat{\gamma} a^r - MNa^{2r-2\mu-2} \quad (r-\mu-1 \leqslant \kappa < r-\mu) \\
&\geqslant a^r \lambda^{r-\mu-2}\left(\hat{\gamma} - \frac{MN^2 a^{r-2\mu-2}}{\lambda^{r-\mu-2}}\right).
\end{aligned}$$

Since $\lambda > a$, there is a number r_0 such that $\left(\hat{\gamma} - \frac{pN^2 a^{r-2\mu-2}}{\lambda^{r-\mu-2}}\right) > 0$, for $r > r_0$. Thus,

$$\frac{\log \mathcal{N}^*(r)}{r \log a} \geqslant \frac{r \log a}{r \log a} + \frac{(r-\mu-2)\log \lambda}{r \log a} + \frac{\log\left(\hat{\gamma} - \frac{MN^2 a^{r-2\mu-2}}{\lambda^{r-\mu-2}}\right)}{r \log a} \quad \text{for } r > r_0,$$

which means that $\underline{D}(A) \geqslant 1 + \log_a \lambda$. So we have that $D(A) \geqslant 1 + \log_a \lambda$. This completes the proof. □

Remark 1. Let the RBIFS and matrix S be as defined above. Suppose that the RBIFSs attractor A is the graph of a continuous function and the connection matrix C is a non-negative matrix which has a strictly positive eigenvector X. Then the box-counting dimension is given by

$$D(A) = \begin{cases} 1 + \log_a \lambda & \text{if } \lambda > a, \\ 2 & \text{otherwise,} \end{cases}$$

where $\lambda = \rho(SC)$.

If there exists a strictly positive eigenvector, it corresponds to the spectral radius of the matrix which is now strictly positive too (see [7,10, p. 551]). Thus, the strictly positive eigenvector X of SC corresponds to the eigenvalue $\lambda = \rho(SC) > 0$ (any other strictly positive eigenvector will also correspond to λ). We can choose $\hat{U}$ and U^* in relations (9) and (10) accordingly and the proof follows that of the theorem. Conditions for a non-negative matrix to have a positive eigenvector are given in [7].

Definition 5. Let $B \geqslant 0$, $B \neq 0$ be a square matrix ($N \times N$). Also, let $x \in \mathbb{R}^N$ be such that $x \geqslant 0$. We define

$$\lambda_1(x) = \max\{\rho : Bx \geqslant \rho x\} \geqslant 0 \quad \text{and} \quad \lambda_2(x) = \min\{\sigma : Bx \leqslant \sigma x\}.$$

We also define

$$\lambda_1(B) = \min\{\lambda_1(x) : x \geqslant 0\} \geqslant 0 \quad \text{and} \quad \lambda_2(B) = \max\{\lambda_2(x) : x \geqslant 0\}.$$

Remark 2. Let the RBIFS be as defined in Theorem 1, but now suppose that C is a reducible matrix and the attractor G of the RBIFS is the graph of a continuous function. Then

if $\lambda_1(SC) > a$, then $1 + \log_a \lambda_1(SC) \leqslant \underline{D}(G) \leqslant \overline{D}(G) \leqslant 1 + \log_a \lambda_2(SC)$,
if $\lambda_2(SC) \leqslant a$, then $D(G) = 2$.

Proof. The proof is similar to the one given for Theorem 1. Suppose that $\lambda_1(SC) > a$ (then $\lambda_2(SC) \geqslant \lambda_1(SC) > a$). From (11) we have

$$\begin{aligned}\mathcal{N}^*(r) &\leqslant \Omega\left((SC)^{\kappa-1}Ra^r + (SC)^{\kappa-1}B\frac{a^{r+1}}{N} + (SC)^{\kappa-2}B\frac{a^{r+2}}{N} + \cdots + (SC)B\frac{a^{r+\kappa-1}}{N}\right.\\ &\quad \left. + B\frac{a^{r+\kappa}}{N} + a^{2\kappa-2}\mathbb{I}\right)\cdot(a+1)^2\\ &= \Omega(\lambda_2(SC)^{\kappa-1}Ra^r + \lambda_2(SC)^{\kappa-1}Ba^{r+1} + \lambda_2(SC)^{\kappa-2}Ba^{r+2}\\ &\quad + \cdots + \lambda_2(SC)Ba^{r+\kappa-1} + Ba^{r+\kappa} + a^{2\kappa-2}\mathbb{I})\cdot(a+1)^2.\end{aligned}$$

Similarly, we obtain $\overline{D}(G) \leqslant 1 + \log_a \lambda_2(G)$.

From (8) we have

$$\begin{aligned}\mathcal{N}^*(r) &\geqslant \Omega\left((SC)^{\kappa-1}\cdot H_1\cdot a^r\right) - MNa^{2\kappa-2}\\ &\geqslant \Omega\left((\lambda_1(SC))^{\kappa-1}\cdot H_1\cdot a^r\right) - MNa^{2\kappa-2}.\end{aligned}$$

Similarly, we obtain $1 + \log_a \lambda_1(SC) \leqslant \underline{D}(G)$. If $\lambda_2(SC) \leqslant a$, we can easily deduce that $\underline{D}(G) = \overline{D}(G) = D(G) = 2$. □

3.1. Special cases

Case I: The bivariate FIS. The bivariate FIS (see [6]) is a RBFIS with connection matrix

$$C_{ij} = 1, \quad i = 1, \ldots, N, \quad j = 1, \ldots, M,$$

$\delta = \frac{1}{N}$, $\psi = 1$ and thus $a = N$. In this case, C is an irreducible non-negative matrix and the spectral radius $\lambda = \rho(SC)$ is given by

$$\lambda = \sum_{i=1}^{N}\sum_{j=1}^{M} |s_{ij}|.$$

Since it has been proved [6] that the attractor of this IFS is a continuous function, from the above theorem we deduce that the box-counting dimension of the attractor A of the bivariate FIS is given by

$$D(A) = \begin{cases} 1 + \log_N \sum_{i=1}^{N} \sum_{j=1}^{M} |s_{ij}| & \text{if } \sum_{i=1}^{N} \sum_{j=1}^{M} |s_{ij}| > N, \\ 2 & \text{otherwise.} \end{cases}$$

Case II: SC is a multiple of a stochastic matrix. Let the RBIFS be as defined in Theorem 1, but now C be reducible. In order for the matrix SC to be a multiple of a stochastic matrix (i.e. the sum of the elements of each row be a constant β) we need:

$$|s_{ij}| = \frac{\beta}{a^2} \quad \text{for } i = 1, \ldots, N, \quad j = 1, \ldots, M.$$

In this case, the spectral radius of SC is $\lambda = \rho(SC) = \beta$ (see [7, p. 84]) and a corresponding eigenvector is $(1, 1, \ldots, 1)^{\mathrm{T}}$. If the attractor A of this BRIFS is a continuous function (see Proposition 1), then the box-counting dimension of A is

$$D(A) = \begin{cases} 1 + \log_a \beta & \text{if } \beta > a, \\ 2 & \text{otherwise.} \end{cases}$$

4. Construction of recurrent bivariate fractal interpolation surfaces

We stated in Theorem 1 that if the attractor of a RFIS is the graph of a continuous function, we can compute its box-counting dimension. Some sufficient assumptions will be given so that the above can be fulfilled.

Proposition 1. *With the same notation as in Section* 2, *assume that for every interval* J_{kl}, $k = 1, 2, \ldots, K$, $l = 1, 2, \ldots, L$, *the points of each of the sets*

$$\{(x_{(k-1)a} = \hat{x}_{k-1},\ y_{(l-1)a+v},\ z_{(k-1)a,(l-1)a+v});\ v = 0, 1, 2, \ldots, a\},$$
$$\{(x_{ka} = \hat{x}_k,\ y_{(l-1)a+v},\ z_{ka,(l-1)a+v});\ v = 0, 1, 2, \ldots, a\},$$
$$\{(x_{(k-1)a+v},\ y_{(l-1)a} = y_{l-1},\ z_{(k-1)a+v,(l-1)a});\ v = 0, 1, 2, \ldots, a\},$$
$$\{(x_{(k-1)a+v},\ y_{la} = y_l,\ z_{(k-1)a+v,la});\ v = 0, 1, 2, \ldots, a\}$$

are collinear. Then there exists a continuous function $f : [0, 1] \times [0, p] \to \mathbb{R}$ *that interpolates the given data* $P = \{(x_i, y_j, z_{ij}) : i = 1, 2, \ldots, N, j = 1, 2, \ldots, M\}$ *and its graph* $\{(x, y, f(x, y)) : (x, y) \in [0, 1] \times [0, p]\}$ *is the attractor A of the RBIFS.*

In [6], Dalla proved that the attractor of a bivariate IFS (BIFS) on $[\alpha, \beta] \times [\gamma, \delta] \times \mathbb{R}$ is the graph of a continuous function that interpolates the data, if the interpolation points are such that the points of each one of the sets

$$\{(x_0,\ y_\mu,\ z_{0,\mu}) : \mu = 0, 1, 2, \ldots, M\},$$
$$\{(x_N,\ y_\mu,\ z_{N,\mu}) : \mu = 0, 1, 2, \ldots, M\},$$
$$\{(x_v,\ y_0,\ z_{v,0}) : v = 0, 1, 2, \ldots, N\},$$
$$\{(x_v,\ y_M,\ z_{v,M}) : v = 0, 1, 2, \ldots, N\}$$

are collinear. The proof of the above proposition is similar.

More generally, we have the following. (The proof is similar to the one presented in [6].)

Proposition 2. *Let the RBIFS be as defined above. Consider the interval* J_{kl}, $k = 1, \ldots, K$, $l = 1, \ldots, L$, *and let, for* $\kappa \in \mathbb{N}$,

$\mathcal{L}^{\kappa}_{kl}[v]$, $v = 1, \ldots, a^{\kappa+1} - 1$,
denote the vertical distance of each one of the points computed in the step κ *of the construction, with* $x = \hat{x}_{k-1}$ *and* $y \in [\hat{y}_{l-1}, \hat{y}_l]$, *from the line defined by the points* $(x_{(k-1)a}, y_{(l-1)a}, z_{(k-1)a,\,(l-1)a})$ *and* $(x_{(k-1)a}, y_{la}, z_{(k-1)a,\,la})$,

$\mathcal{R}^{\kappa}_{kl}[v]$, $v = 1, \ldots, a^{\kappa+1} - 1$,
denote the vertical distance of each one of the points computed in the step κ *of the construction, with* $x = \hat{x}_k$ *and* $y \in [\hat{y}_{l-1}, \hat{y}_l]$, *from the line defined by the points* $(x_{ka}, y_{(l-1)a}, z_{ka,\,(l-1)a})$ *and* $(x_{ka}, y_{la}, z_{ka,\,la})$,

$\mathcal{D}^{\kappa}_{kl}[v]$, $v = 1, \ldots, a^{\kappa+1} - 1$,
denote the vertical distance of each one of the points computed in the step κ *of the construction with* $y = \hat{y}_{l-1}$ *and* $x \in [\hat{x}_{k-1}, \hat{x}_k]$, *from the line defined by the points* $(x_{(k-1)a}, y_{(l-1)a}, z_{(k-1)a,\,(l-1)a})$ *and* $(x_{ka}, y_{(l-1)a}, z_{ka,\,(l-1)a})$,

$\mathcal{U}^{\kappa}_{kl}[v]$, $v = 1, \ldots, a^{\kappa+1} - 1$,
denote the vertical distance of each one of the points computed in the step κ *of the construction with* $y = \hat{y}_l$ *and* $x \in [\hat{x}_{k-1}, \hat{x}_k]$, *from the line defined by the points* $(x_{(k-1)a}, y_{la}, z_{(k-1)a,\,la})$ *and* $(x_{ka}, y_{la}, z_{ka,\,la})$.

Each one of these vertical distances is taken as positive if the corresponding interpolation point is above the corresponding line; otherwise it is taken as negative. If $\kappa = 0$, *then* $\mathcal{L}^{0}_{kl}[v]$ *denotes the vertical distance of the interpolation points from the straight line defined above, etc. If we can select the vertical scaling factors so that*

$$s_{i,j} \cdot \mathcal{R}^{\kappa}_{\mathbb{J}(i,j)}[v] = s_{i+1,j} \cdot \mathcal{L}^{\kappa}_{\mathbb{J}(i+1,j)}[v],$$

$$s_{i,j} \cdot \mathcal{U}^{\kappa}_{\mathbb{J}(i,j)}[v] = s_{i,j+1} \cdot \mathcal{D}^{\kappa}_{\mathbb{J}(i,j+1)}[v],$$

$$\text{for } i, j, v \in \mathbb{N} : \ i = 1, \ldots, N-1, \ j = 1, \ldots, M-1, \ v = 1, \ldots, a^{\kappa+1} - 1, \kappa \in \mathbb{N},$$

then there exists a continuous function $f : [0, 1] \times [0, p] \to \mathbb{R}$ *that interpolates the given data* $\Delta = \{(x_i, y_j, z_{ij}) : i = 1, 2, \ldots, N;\ j = 1, 2, \ldots, M\}$ *and its graph* $\{(x, y, f(x, y)) : (x, y) \in [0, 1] \times [0, p]\} = A$.

Of course in general, it is extremely difficult to find a RBIFS satisfying the above conditions. A special case is described below.

Corollary 1. *Let the RBIFS be as defined above, and assume that*

$$\mathcal{R}^{0}_{\mathbb{J}(i,j)}[v] = \mathcal{L}^{0}_{\mathbb{J}(i,j)}[v],$$

$$\mathcal{U}^{0}_{\mathbb{J}(i,j)}[v] = \mathcal{D}^{0}_{\mathbb{J}(i,j)}[v],$$

for $v = 1, 2, \ldots, a-1$ *and* $s_{ij} = s < 1$, *for* $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, M$. *Then the attractor of the above RBIFS is the graph of a continuous function that interpolates the given data* Δ.

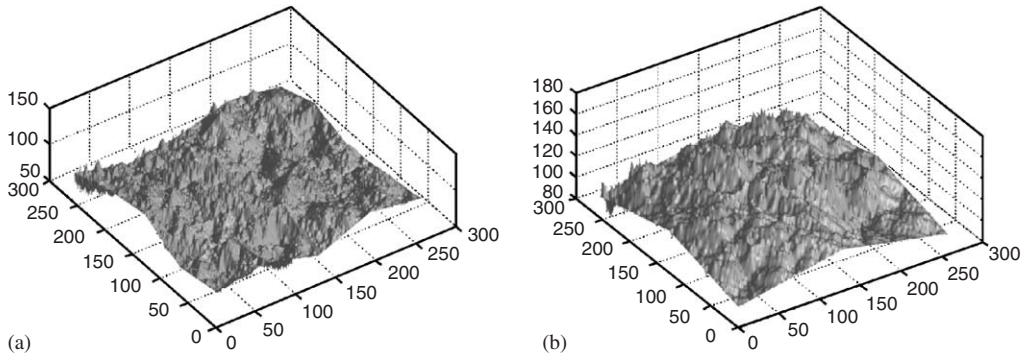

Fig. 1. Two fractal interpolation surfaces constructed according to the conditions described in Proposition 1. The first (a) (where $N = M = 8$, $K = L = 4$) has box-counting dimension 2.2769 and the second (b) (where $N = M = 4$, $K = L = 2$) 2.3325.

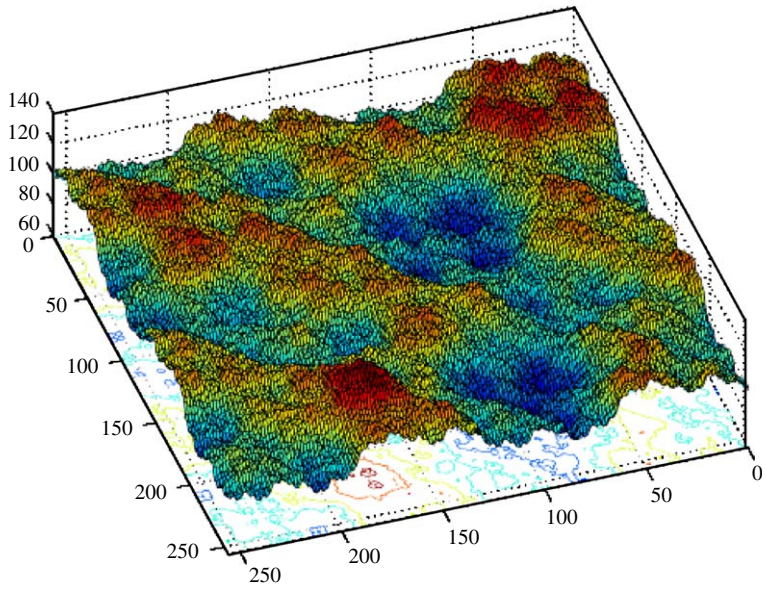

Fig. 2. The RIFS (defined on $[0, 256] \times [0, 256] \times \mathbb{R}$) satisfies the conditions in Corollary 1. Hence its attractor is the graph of a continuous function. Here $M = N = 6$, $K = L = 2$, $a = 3$. The connection vector is $C^v = (1, 3, 4, 1, 2, 2, 1, 3, 3, 4, 1, 1, 2, 3, 1, 4, 3, 1, 2, 3, 4, 1, 4, 2, 2, 4, 3, 1, 2, 1, 3, 4, 2, 4, 1, 2)$. The interpolation points are shown in Table 1 and $s = 0.38$. The matrix SC is irreducible and the box-counting dimension of the attractor is 2.3319.

4.1. Examples

In Figs. 1 and 2, the attractors of some RBIFSs on $\mathbb{R}^3$ are shown. The chosen RBIFSs satisfy either Proposition 1 or Corollary 1.

5. Approximation of natural surfaces using RBIFSs

One may use RBIFSs to approximate any discrete natural surface, as complicated as may be, using the methodology described below. This methodology is based on ideas similar to the ones

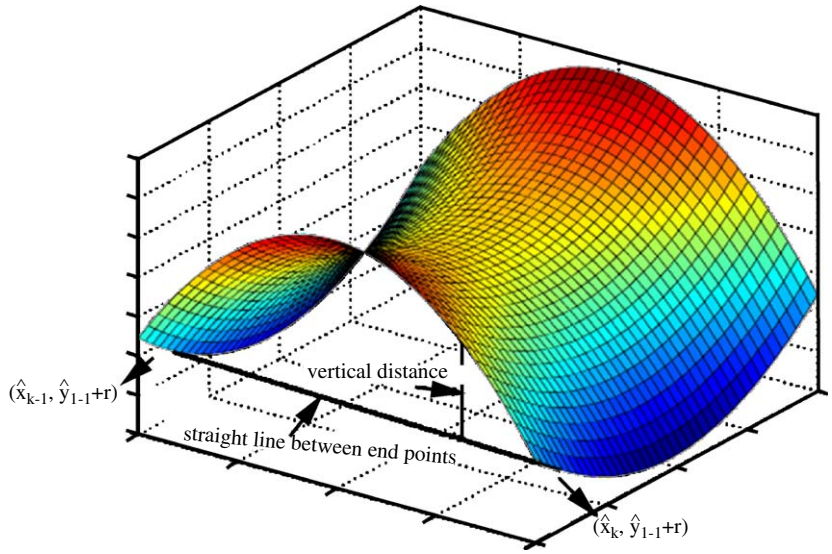

Fig. 3. Computation of the contractivity factors.

presented by Mazel and Hayes in [15], where affine FIFs where used to model single-valued discrete sequences.

Proposition 2 gives the general conditions that the interpolation points and the contractivity factors must satisfy so that the attractor of the corresponding RIFS be a continuous function. If the interpolation points are collinear on the boundary of each interval J_{kl}, then any selection of the contractivity factors s_{ij} will be sufficient (Proposition 1). On the other hand, if the interpolation points are not collinear on the boundary of each interval J_{kl}, then it is almost impossible to find contractivity factors that satisfy the conditions of Proposition 2 for an arbitrary selection of interpolation points. However, for $\kappa = 0$ it is relatively easy to find contractivity factors such that

$$s_{i,j} \cdot \mathcal{R}^0_{\mathbb{J}(i,j)}[v] = s_{i+1,j} \cdot \mathcal{L}^0_{\mathbb{J}(i+1,j)}[v],$$

$$s_{i,j} \cdot \mathcal{U}^0_{\mathbb{J}(i,j)}[v] = s_{i,j+1} \cdot \mathcal{D}^0_{\mathbb{J}(i,j+1)}[v],$$

for $i = 1, \ldots, N-1$, $j = 1, \ldots, M-1$, $v = 1, \ldots, a-1$, or

$$|s_{i,j} \cdot \mathcal{R}^0_{\mathbb{J}(i,j)}[v] - s_{i+1,j} \cdot \mathcal{L}^0_{\mathbb{J}(i+1,j)}[v]| < \varepsilon, \tag{15}$$

$$|s_{i,j} \cdot \mathcal{U}^0_{\mathbb{J}(i,j)}[v] - s_{i,j+1} \cdot \mathcal{D}^0_{\mathbb{J}(i,j+1)}[v]| < \varepsilon, \tag{16}$$

for $\varepsilon > 0$ (relatively small). In this case, the attractor of the corresponding RIFS is not graph of a continuous surface. Instead, the attractor is a compact subset of $[0, 1] \times [0, p] \times \mathbb{R}$ that approximates a continuous surface.

Now, consider the data set $D = \{(n, m, f(n, m)) : n = 0, 1, \ldots, N^*; m = 0, 1, \ldots, M^*\}$ representing points of an arbitrary surface. Our goal is to choose interpolation points and contractivity factors such that the attractor of the corresponding RIFS approximates the surface. We choose $\delta, \psi \in \mathbb{N}$ a priori and form the sections I_{ij} and the intervals J_{kl}, so that each section contains $(\delta + 1) \times (\delta + 1)$ points of D and each interval $(\psi + 1) \times (\psi + 1)$ points of D. For

Fig. 4. Any grey-scale image can be viewed as a (discrete) surface. In (a), the test image of Lena is shown. Each attractor of the RIFSs that approximates the "Lena" surface (for different values of δ, ψ and error tolerance) is shown in (b)–(d).

each section I_{ij} we seek the best-mapped interval J_{kl} with respect to a metric h, using bivariate mappings. We compute the contractivity factor of that mapping as follows.

Recall that bivariate mappings have the property to map vertical lines (parallel to zz') to vertical lines scaled by the vertical scaling factor s. Let $|\mu_r^y|$, $r = 0, 1, \ldots, \psi$, be the mean absolute vertical distance between any value of $f(x, y)$ lying inside the interval J_{kl} (where $x = \hat{x}_{k-1}, \hat{x}_{k-1} + 1, \ldots, \hat{x}_{k-1}+\psi = \hat{x}_k, y = \hat{y}_{l-1}+r$) and the straight line between the vertices $(\hat{x}_{k-1}, y)$ and $(\hat{x}_k, y)$ (see Fig. 3). The sign of μ_r^y is taken as positive, if $f(x, y)$ is above the straight line and negative otherwise. Similarly, let $|\mu_r^x|$, $r = 0, 1, \ldots, \psi$, be the mean absolute vertical distance between any function value $f(x, y)$ (where $y = \hat{y}_{l-1}, \hat{y}_{l-1}+1, \ldots, \hat{y}_{l-1}+\psi = \hat{y}_l, x = \hat{x}_{k-1}+r$) and the straight line between the vertices $(x, \hat{y}_{l-1})$ and $(x, \hat{y}_l)$. Let $\mu = \text{mean}\{\mu_r^x, \mu_r^y, r = 0, 1, \ldots, \psi\}$.

Table 1
The interpolation points used for the RIFS shown in Fig. 2

	Δ						
y	x						
	0	$\frac{256}{6}$	$2\frac{256}{6}$	128	$4\frac{256}{6}$	$5\frac{256}{6}$	256
0	100	115	090	100	115	090	100
$\frac{256}{6}$	115	122	110	115	090	110	115
$2\frac{256}{6}$	090	110	080	090	110	120	090
128	100	115	090	100	115	090	100
$4\frac{256}{6}$	115	090	100	115	090	100	115
$5\frac{256}{6}$	090	110	080	090	125	110	090
256	100	115	090	100	115	090	100

We compute μ^* similarly by using the values of the function lying inside I_{ij}. Then, the contractivity factor is given by the ratio μ^*/μ.

Since the contractivity factor has been chosen, the remaining parameters of the bivariate map w are computed by Eqs. (5) and the set $w(\{(x, y, f(x, y)) : (x, y) \in J_{kl}\}$ is formed. Then we compute the distance between the sets $\{(x, y, f(x, y)) : (x, y) \in I_{i,j}\}$ and $w(\{(x, y, f(x, y)) : (x, y) \in J_{kl}\}$ and repeat the procedure for every interval. If the contractivity factor that has been computed does not satisfy conditions (15)–(16) or is greater than 1, we remove the corresponding interval from the search pool. Finally, we choose the interval that minimises the previously mentioned distance. If this distance, however, is greater than an error tolerance value (chosen a priori) we split the section to four subsections (adding new interpolation points) and repeat the procedure for each new subsection.

The collage theorem for RIFSs (see [3]) ensures that the attractor of the emerging RIFS will approximate the original surface f. The methodology is described in detail in [4], where it is successfully used to model and compress grey-scale images (see Fig. 4). The compression is achieved by storing only the map parameters of the RIFS (interpolation points, contractivity factors and connection vector) instead of all the pixel values (i.e. the set D) of the image. The examination of the conditions (15)–(16) for each selected interval J_{kl} and corresponding contractivity factor significantly improves the speed of the procedure, as a lot of the intervals are removed from the search pool. In addition, the use of bivariate mappings instead of affine ones improves the quality of the reconstructed surface. In [4], the proposed fractal interpolation approach is compared to the previously presented fractal methods and found to give more satisfactory results.

References

[1] M.F. Barnsley, Fractal functions and interpolation, Constr. Approx. 2 (1986) 303–329.
[2] M.F. Barnsley, Fractals Everywhere, second ed., Academic Press Professional, New York, 1993.
[3] M.F. Barnsley, J.H. Elton, D.P. Hardin, Recurrent iterated function systems, Constr. Approx. 5 (1989) 3–31.
[4] P. Bouboulis, Leoni Dalla, V. Drakopoulos, Image compression using recurrent bivariate fractal interpolation surfaces, Internat. J. Bifur. Chaos (2006) (to appear).
[5] B.A. Cambell, M.K. Shepard, Shadows on a planetary surface and implications for photometric roughness, ICARUS 134 (1998) 279–291.

[6] L. Dalla, Bivariate fractal interpolation functions on grids, Fractals 10 (1) (2002) 53–58.
[7] F.R. Gantmacher. Matrix Theory, vol. 2, Chelsea Publishing Company, New York, 2000.
[8] J.S. Geronimo, D. Hardin, Fractal interpolation surfaces and a related 2d multiresolutional analysis, J. Math. Anal. Appl. 176 (1993) 561–586.
[9] D.P. Hardin, P.R. Massopust, Fractal interpolation functions from $\mathbb{R}^n \to \mathbb{R}^m$ and their projections, Z. Anal. Anw. 12 (1993) 535–548.
[10] S. Karlin, H.M. Taylor, A First Course in Stochastic Processes, second ed., Academic Press, New York, 1974.
[11] R. Malysz, The Minkowski dimension of the bivariate fractal interpolation surfaces, Chaos, Solitons & Fractals 27 (5) (2006) 1147–1156.
[12] B.B. Mandelbrot, D.E. Passoja, A.J. Paullay, Fractal character of fracture surfaces of metals, Nature 308 (1984) 721–722.
[13] P.R. Massopust, Fractal surfaces, J. Math. Anal. Appl. 151 (1) (1990) 275–290.
[14] D.S. Mazel, Representation of discrete sequences with three-dimensional iterated function systems, IEEE Trans. Signal Process. 42 (1994) 3269–3271.
[15] D.S. Mazel, M.H. Hayes, Using iterated function systems to model discrete sequences, IEEE Trans. Signal Process. 40 (1992) 1724–1734.
[16] B.B. Nakos, C. Mitsakaki, On the fractal character of rock surfaces, Internat. J. Rock. Mech. Min. Sci. Geomech. Abstr. 28 (1991) 527–533.
[17] M.A. Navascues, M.V. Sebastian, Generalization of Hermite I. functions by fractal interpolation, J. Approx. Theory 131 (2004) 19–29.
[18] C.S. Pande, L.E. Richards, N. Louat, B.D. Dempsey, A.J. Schwoeble, Fractal characterization of fractured surfaces, Acta Metalurgica 35 (1987) 1633–1637.
[19] J.R. Price, Resampling and reconstruction with fractal interpolation functions, IEEE Signal Process. Lett. 5 (1998) 228–230.
[20] P. Wong, J. Howard, J. Li, Surfaces roughening and the fractal nature of rocks, Phys. Rev. Lett. 57 (1986) 637–640.
[21] H. Xie, H. Sun, The study of bivariate fractal interpolation functions and creation of fractal interpolation surfaces, Fractals 5 (4) (1997) 625–634.
[22] H. Xie, H. Sun, Y. Zu, Z. Feng, Study on generation of rock fracture surfaces by using fractal interpolation, Internat. J. Solids Struct. 38 (2001) 5765–5787.
[23] N. Yokoya, K. Yamamoto, N. Funakubo, Fractal-based analysis and interpolation of 3d natural surface shapes and their application to terrain modeling, Comput. Vision Graphics Image Process. 46 (1989) 284–302.
[24] N. Zhao, Construction and application of fractal interpolation surfaces, Visual Computer 12 (1996) 132–146.