Writing Sequences on the Plane

Emina Soljanin, *Member, IEEE*

Invited Paper

***Abstract*—The problem of arranging two-dimensional arrays of data into one-dimensional sequences comes up in image processing, color quantization, and optical and magnetic data recording. A good arrangement should enable the one-dimensional sequences to be modeled as Markov chains or shifts of finite type. Since this is not possible in general, two-dimensional data is most commonly scanned by rows, columns, or diagonals. We look into three unusual ways to write a sequence in the plane: by Penrose tilings, by space-filling curves, and by cylindrical and spiral lattices. We show how Penrose tilings can be used to record information and how some spiral lattices can be used for quantization of color spaces.**

***Index Terms*—Color quantization, cylindrical and spiral lattices, image processing, information recording, Penrose tilings, space-filling curves.**

I. Introduction

A NUMBER of communications problems require multidimensional arrays of data to be arranged in one-dimensional sequences. The most common example is image processing, where two-dimensional arrays of image data are read in a sequence by scanning the image [1]–[3]. Another example is quantization of three-dimensional color spaces such as RGB or Lab, where color palettes with each color described by a single number makes color-quantized images more amenable to the type of processing used for gray-scale ones [4]–[7].

Traditional storage systems record data in one-dimensional tracks. More advanced optical recording systems store data in two-dimensional arrays, but are capable of a *one-shot* read/write of the entire storage medium. Significant progress has been made in developing the same feature for magnetic-disk storage. This has generated interest in two-dimensional (d, k) and other codes which write data in rectangular grids. The first question to ask about these codes concerns, of course, their maximum rate. While the capacities of one-dimensional (d, k) constraints can be easily computed by linear algebra, the capacities of their two-dimensional counterparts are known only in some cases. There are, however, numerous results on bounding capacities of two-dimensional (d, k) constraints and characterizing of zero capacity constraints, see [8]–[14] and references therein. Many of these results are based on combinatorial ideas first used to bound the asymptotic growth rate of the number of independent sets in a grid graph [15]. This problem is in particular related to the two-dimensional $(1, \infty)$ constraint consisting of two-dimensional arrays in which each 1 has a 0 to the left, right, above, and below. The bounds in [15] are derived by reducing the problem to one dimension, namely, by considering a two-dimensional array as an infinite sequence of finite-length columns and then letting the column length go to infinity. Similarly, one can scan two-dimensional grids over rows or diagonals. A diagonal scan was used in [9] to find bit-stuffing bounds on the capacity of two-dimensional constrained arrays.

Whenever two-dimensional data with spatial dependencies are to be arranged in a temporal sequence for analysis or processing, the problem reduces to writing a sequence of natural numbers on the plane so that points close on the plane are labeled by numbers which are not very far apart. Such numbering allows the temporal sequence to be modeled as a well-understood object: a Markov chain or a shift of finite type. Clearly, this is not possible by row, column, or diagonal scans of two-dimensional data written in rectangular grids.

We look into three unusual ways to write a sequence in the plane: by Penrose tilings, by space-filling curves, and by cylindrical and spiral lattices. We revisit the topic of space-filling curves because they have been successfully used in image processing (but very few other areas) as special grid scans which visit every cell in the quadrant they enter before exiting the quadrant [1]–[3]. In a special spiral lattice known as the Fibonacci lattice, knowing the time index of a point is sufficient to determine the time indexes of its neighbors since each is only a Fibonacci number apart. We have recently used the Fibonacci lattice for quantization of color spaces and processing of color images [7]. Since Penrose tilings can be parameterized by binary sequences which do not contain consecutive 1's, we were interested to find out what kind of a recording channel they would make [6]. The paper is organized as follows. In Section II, we look at Penrose tilings from a Shannon theory point of view, and examine whether they can be used as a tool for recording information. In Section III, we discuss space-filling curves, and review their use in image processing. In Section IV, we study cylindrical and spiral lattices, and show how they can be used for quantization of color spaces.

II. Shannon Discrete Noiseless Systems and Penrose Tilings

Penrose tilings of the plane have been studied using tools of various areas of mathematics such as algebra, combinatorics,

Manuscript received April 3, 2001; revised December 21, 2001. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Cambridge, MA, August 1998.

The author is with Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: emina@lucent.com).

Communicated by S. Shamai, Guest Editor.

Publisher Item Identifier S 0018-9448(02)04012-9.

0018-9448/02$17.00 © 2002 IEEE



(a) (b)

Fig. 1. Plane tilings by (a) Penrose kites and darts and (b) Penrose rhombuses.

dynamical system theory, theory of quasi-crystals, and even topology and K-theory. We study the tilings as means of recording information. It is known that the uncountable set of all Penrose tilings can be parameterized by binary sequences which do not contain consecutive 1's. It was surprising to find out that the most popular Penrose patterns come from (in a certain sense) nontypical *source* sequences, and that the Penrose tiling *channel* capacity is zero.

A. Penrose Tilings of the Plane

A *plane tiling* $\mathcal{T}$ is a partition of the plane into a countable number of closed sets T_i called *tiles*

$$\mathcal{T} = \{T_1, T_2, \ldots\}$$

where the interiors of the sets T_i are pairwise disjoint and their union is the whole plane. Two plane tilings are shown in Fig. 1. They are known as Penrose tilings by kites and darts shown in Fig. 1(a) and by thick and thin rhombuses shown in Fig. 1(b). Note that in both cases, each set T_i is congruent to one of the two possible tiles. However, the set of all possible tilings is uncountable. Penrose tilings represent interesting objects in various branches of mathematics, which has been discovered by de Bruijn in algebra, Connes in topology and K-theory, Conway in combinatorics, and Robinson in dynamical system theory, see [16]–[21]. We are interested in Penrose tilings as means of communicating information. This information-theoretic study is based on the Robinson analysis of Penrose tilings [19] described below .

A Penrose tiling by kites and darts has a corresponding tiling by two triangular tiles L_A and S_A, shown in Fig. 2(a), obtained by cutting the kites and darts along their axes of symmetry, as indicated in Fig. 1(a). The edge length τ in the figure equals $(1+\sqrt{5})/2$. We define a tiling of type A as a plane tiling by L_A and S_A which satisfies the matching rules specified by the arrows on the tile edges and color (black and white) on their vertices. A triangle tiling of type A has a corresponding tiling by kites and darts, which can be obtained by deleting the triangle edges connecting either two black or two white vertices. This is illustrated in Fig. 1(a), where the "deleted" edges are shown in gray. Similarly, a Penrose tiling by thick and thin rhombuses has a corresponding plane tiling by two triangular tiles L_B and S_B, shown in Fig. 2(b), obtained by cutting the thick rhombuses along their long diagonal and the thin rhombuses along their short diagonal, as indicated in Fig. 1(b). We define a tiling of type B as a plane tiling by L_B and S_B which satisfies the matching rules specified by the arrows on the tile edges and

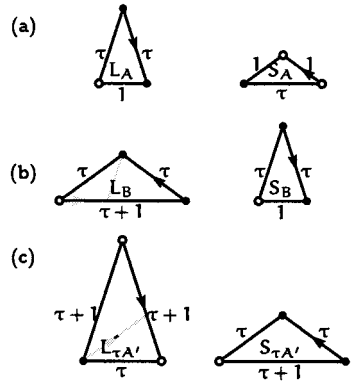


Fig. 2. Elementary triangle tiles.

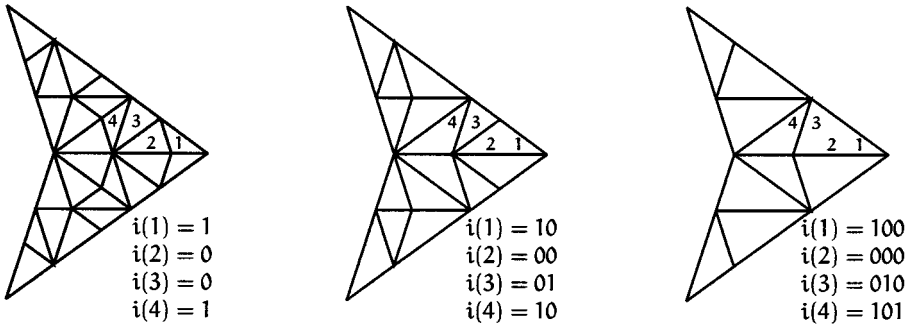
color (black and white) on their vertices. A tiling of type B has a corresponding tiling by thick and thin rhombuses obtained by deleting the long edges of L_B triangles and the short edges of S_B triangles.

Let $\mathcal{T}_0$ be a type A tiling (as in Fig. 3(a)) whose tiles are congruent to triangles L_A and S_A shown in Fig. 2(a). If we delete all the edges in $\mathcal{T}_0$ that join two vertices of different color and separate L_A-tiles from S_A-tiles, we obtain a new tiling of the plane $\mathcal{T}_1$, as in Fig. 2(b). Its tiles are congruent to triangles L_B and S_B (shown in Fig. 2(b)), and therefore $\mathcal{T}_1$ is a type B tiling. If we then delete all the edges in $\mathcal{T}_1$ that join two vertices of the same color and separate L_B-tiles from S_B-tiles, we obtain $\mathcal{T}_2$, a new type A tiling whose tiles are congruent to triangles $L_{\tau A'}$ and $S_{\tau A'}$ shown in Fig. 3(c). By iterating this edge-deleting procedure, we obtain a sequence $\{\mathcal{T}_n\}$ of plane tilings by large triangles L_n and small triangles S_n. Given a tiling $\mathcal{T}_0$, we can associate to each of its tiles T a binary sequence $\{x_n\}$ such that $x_n = 0$ if the triangle of T_n containing T is large and $x_n = 1$ if that triangle is small. We denote this sequence by $i(\mathcal{T}_0, T)$ and refer to it as the index sequence of tile T in tiling $\mathcal{T}$. Since by the edge-deleting procedure, a small triangle S_n *always* becomes a part of a large triangle L_{n+1}, an index sequence does not contain the string 11. Index sequences of four example tiles are shown in Fig. 3.

B. The Golden Mean Shift and Index Sequences for Penrose Tilings

We have seen in the previous subsection that an index sequence $i(\mathcal{T}_0, T)$ of tile T in type A or B tiling $\mathcal{T}$ does not contain string 11. It is also true that every infinite binary sequence $\{x_n\}$ such that

$$x_n = 1 \Longrightarrow x_{n+1} = 0 \tag{1}$$



(a) (b) (c)

Fig. 3. Edge deleting and index sequences.

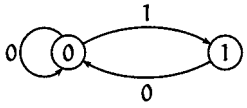


Fig. 4. Graph for the golden mean shift.

is an index sequence $i(\mathcal{T}_0, T)$ of some tile T in some type A or B tiling $\mathcal{T}$ [19]. We denote by X the set of all sequences $\{x_n\}$ that satisfy (1). Set X can be obtained by walks on the graph shown in Fig. 4, which we shall refer to as the *golden mean shift* [22, p. 100]. Each length-n sequence in X corresponds to the sequence of states in an n-step walk in this graph. In the terminology of (d, k) codes, the golden mean shift is the $(1, \infty)$ constrained system. We can analyze different Penrose tilings by analyzing the constrained system X using tools of symbolics dynamics.

Let $B_n(X)$ denote the set of all index sequences of length n. The size of $B_n(X)$, denoted by $|B_n(X)|$, grows exponentially with n. The asymptotic growth rate of $|B_n(X)|$ is called the *topological entropy* of the constraint $h(X)$

$$h(X) = \lim_{n\to\infty} \frac{\log|B_n(X)|}{n}. \tag{2}$$

For the golden mean shift, numbers $|B_n(X)|$ form the familiar Fibonacci sequence $\{F_n\}$ with $|B_n(X)| = F_{n+1}$. The sequence of numbers $F_n, n \geq 0$, defined recursively as follows:

$$\begin{aligned} F_0 &= 1, \qquad F_1 = 1 \\ F_n &= F_{n-1} + F_{n-2}, \qquad n \geq 2 \end{aligned} \tag{3}$$

is commonly known as the Fibonacci sequence. It was first defined in 1202 by the Italian mathematician Leonardo of Pisa, nicknamed Fibonacci, as a solution to his rabbit population growth puzzle. Assume that at breeding season zero, we start with a single pair of immature rabbits, which mature in one season. Every season, each mature pair produces precisely one immature pair, which again takes one season to mature. Under this assumptions, if rabbits are immortal, the number of pairs of rabbits in each season F_n satisfy recursion (3). The assumptions, of course, do not hold in real life. But Fibonacci numbers really occur in nature, as we shall see and use in Section IV-A.

The ratios of consecutive Fibonacci numbers converge to the *golden mean*

$$\tau = \lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \frac{\sqrt{5}+1}{2}$$

also referred to as the golden number, section, or ratio. The topological entropy of the golden mean shift equals $\log\tau$. Hence the name of the shift.

The golden mean is considered to be one of nature's special numbers (like π or e), particularly related to the geometry of plants and shells. More on geometric and numerical patterns in plants can be found in [23], and more technical references therein such as [24]. The golden number was known to the ancient Greeks, who thought that a rectangle with the ratio of the sides equal to the golden number has the most pleasing proportions. Such a rectangle can be subdivided into a square and a smaller rectangle that has the same proportions as the larger one.

Several tilings by Penrose kites and darts have received special attention, often because of the symmetry they possess. They are drawn in Fig. 5, by supplying only a single sequence to the drawing program. The *infinite sun pattern* is a tiling which has pentagonal symmetry and five kites at its center. The index sequence of any tile in the infinite sun pattern ultimately agrees with

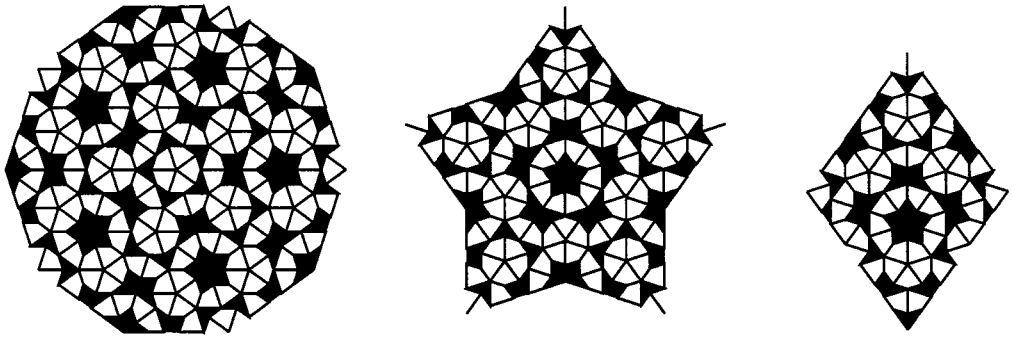
$$0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, \ldots .$$

The *infinite star pattern* is a tiling which has pentagonal symmetry and five darts at its center. The index sequence of any tile in the infinite star pattern ultimately agrees with

$$1, 0, 0, 0, 1, 0, 0, 0, 1, 0, \ldots .$$

The *cartwheel* is a tiling by Penrose kites and darts which has been described by Gardner as the most extraordinary of all Penrose tilings, essential for understanding these tilings in general [20]. It interesting to note that the index sequence of any tile in the cartwheel ultimately agrees with the all-zero sequence.

Most Penrose tilings, however, do not have any symmetries. To see that, we look into a *universally large* set of sequences in X called *doubly transitive* sequences. A sequence $\{x_n\}$ in a shift space X is *doubly transitive* if every block in X appears in x infinitely often to the left and to the right. It is intuitively clear that any such sequence *tells more* about the shift X than,



(a) (b) (c)

Fig. 5. (a) Sun, (b) star, and (c) cartwheel patterns.

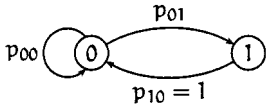


Fig. 6. $(1, \infty)$ constraint.

say, the all-zero sequence. It can be shown that the set of all doubly transitive sequences in X is the complement of a null set with respect to any stationary Markov chain on the golden mean graph in Fig. 4 which assigns positive conditional probability to every edge of the graph [22, p. 334]. (This holds for a large class of shift spaces.) Periodic sequences (such as the underlying index sequences of the sun, star, and cartwheel patterns) are not doubly transitive for the golden mean shift and therefore belong to a set of measure zero for any Markov chain as defined above.

C. The Source

Can Penrose tilings be used to record sequences on the plane? An infinite sequence generated by any binary source satisfying the constraint (1) can be *recorded* on the plane as a Penrose tiling. Our source is any stationary Markov chain on the graph shown in Fig. 4 which assigns positive conditional probability to every edge in the graph. The maximum entropy rate of such a source is equal to $\log \tau$, and is achieved by a unique probability assignment: $p_{00} = 1/\tau$ and $p_{01} = 1/\tau^2$.

D. The Channel

We illustrate the process of recording of binary sequences from the source in Fig. 6 by two examples. The first, shown in Fig. 7, explains the recording of sequence $(0, 0, 0, 0, 0, \ldots)$. We start with a type A tiling, and therefore the first bit 0 is recorded as an L_A triangle. The second bit 0 tells that the previous triangle is a part of a L_B triangle. Recall that the way to draw L_B is determined by the matching rules. The last two 0's are recorded in the same manner. We can see that recording of an infinite all-zero sequence would result in a half-plane tiling which can then be reflected to obtain a tiling of the plane. The corresponding Penrose tiling is the cartwheel.

The second example, shown in Fig. 8, explains the recording of sequence $(0, 0, 1, 0, 0, \ldots)$. We start with a type A tiling, and, therefore, the first bit 0 is recorded as an L_A triangle. The second bit 0 tells that the previous triangle is a part of an L_B triangle. The third bit 1 tells that the previous triangle is an $S_{\tau A'}$ triangle. We can see that recording of an infinite sequence which periodically repeats the last four-bit pattern would result in a tiling of the cone defined by the equal-length edges of the first tile. A tiling of the plane can then be obtained by reflections. The corresponding Penrose tiling is the infinite sun pattern.

Note that in both examples, the recording of each 0 after the third bit results in the addition of a larger and larger unpartitioned triangle to the area used for recording of the sequence up to that point. To obtain a Penrose tiling with L_A and S_A tiles, we partition each triangle larger than L_A and S_A into two triangles, repeating this process until all tiles are congruent to either L_A or S_A. Since this partitioning must not violate the matching rules, it is unique and therefore carries no additional information. The impact of this on the channel capacity is explained below.

In his study of discrete, noiseless transmission channels Shannon defined the channel capacity C as

$$C = \lim_{T \to \infty} \frac{\log N(T)}{T}$$

where $N(T)$ is the number of allowed signals of duration T [25]. (Note that the capacity is equal to the topological entropy defined by (2) when each symbol has the same duration.) The plane is a recording channel whose resource is area rather than time and information is communicated by patterns rather than sequences. The channel capacity must therefore be defined as

$$C = \lim_{A \to \infty} \frac{\log N(A)}{A} \tag{4}$$

where $N(A)$ is the number of allowed patterns of area A. To compute this capacity, we represent the plane as a discrete, constrained communications channel shown in Fig. 9. The channel state labels tell whether the used up area is a large L or a small S triangle. The graph edge labels refer to the cost of using the channel: they tell how many times the area increases with the transition. From Fig. 9, we see that the capacity of the Penrose tiling channel is given by

$$C = \lim_{n \to \infty} \frac{\log N(\tau^n A_0)}{\tau^n A_0}$$

where A_0 is the area of L_A and $N(\tau^n A_0)$, the number of allowed patterns of area $\tau^n A_0$, equals to the number of sequences of length n satisfying the $(1, \infty)$ constraint. Therefore,

$$C = \lim_{n \to \infty} \frac{n \log \tau}{\tau^n A_0} = 0.$$

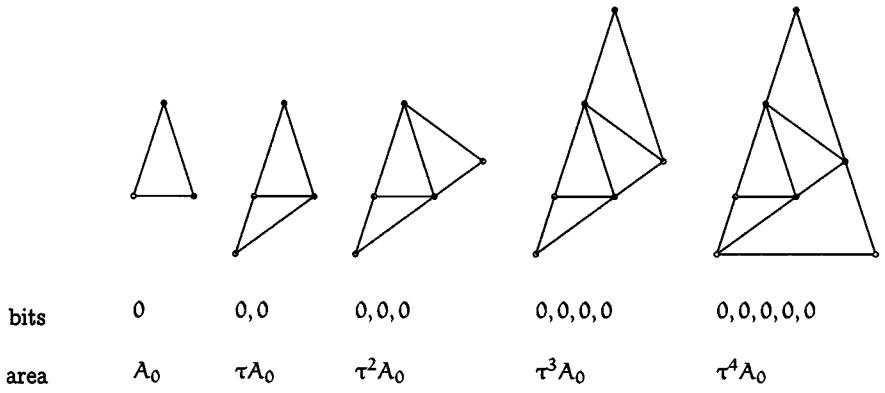


Fig. 7. Recording of sequence $(0, 0, 0, 0, 0, \ldots)$.

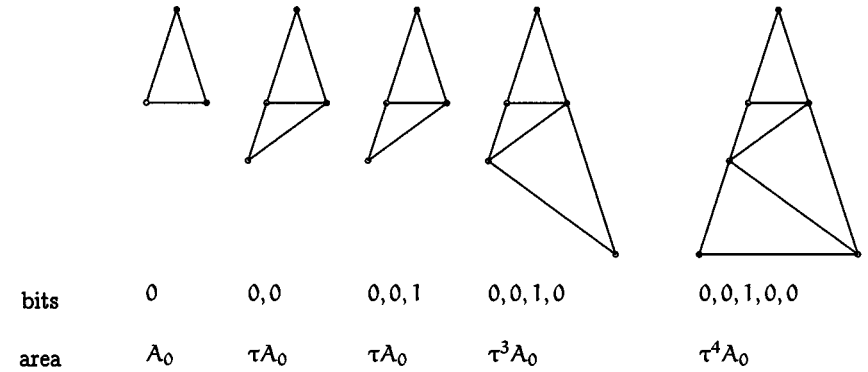


Fig. 8. Recording of sequence $(0, 0, 1, 0, 0, \ldots)$.

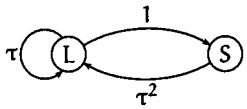


Fig. 9. Communications channel representation of Penrose tilings.

Several different types of tiling go under the name of generalized Penrose tiling, and are studied by the theory of quasi-crystals [21]. In one simple generalization, the partitioning of large triangles is at each level decided by a coin flip rather than by the matching rules [21]. It would be interesting to study the impact of this change to the tiling channel capacity.

III. Space-Filling Curves

Continuous curves that pass through every point of the square or cube are known as space-filling curves. In 1878, Cantor showed that the interval $[0, 1]$ can be mapped bijectively onto the square $[0, 1]^2$. That was a consequence of his more general result that any two finite-dimensional smooth manifolds, no matter what their dimensions, have the same cardinality. In 1879, Netto demonstrated that a bijective map from $[0, 1]$ to $[0, 1]^2$ is necessarily discontinuous. Dropping the requirement that the mapping be one-to-one, in 1890, Peano found a continuous surjective map from the interval onto the square, i.e., constructed the first space-filling curve. Further examples by Hilbert, Moore, Lebesgue, Sierpiński, Pólya, and others followed. A comprehensive treatment of the subject can be found in the book by Sagan [26].

Hilbert proposed an illuminating, geometric way to construct space-filling curves, based on the following principle. If the interval $\mathcal{I}$ can be mapped continuously onto the square $\mathcal{Q}$, then after partitioning $\mathcal{I}$ into four congruent subintervals and $\mathcal{Q}$ into four congruent subsquares, each subinterval can be mapped continuously onto one of the subsquares. Starting with such partitioning and repeating the process n times, $\mathcal{I}$ and $\mathcal{Q}$ are partitioned into 2^{2n} congruent replicas. Hilbert has demonstrated that the subsquares can be arranged so that adjacent subintervals correspond to adjacent subsquares with an edge in common. Therefore, if a square corresponds to an interval, then its subsquares correspond to subintervals of that interval. Fig. 10 illustrates how this process is carried out for the first four steps for what is known as the Hilbert space-filling curve. The polygonal line indicates the order in which the subsquares are to be taken so that if a square corresponds to an interval, then its subsquares correspond to subintervals of that interval.

The most intensive mathematical research activities on the subject of space-filling curves took place around the turn of the century. The interest in the subject was refreshed with the relatively recent interest in fractals. In the area of telecommunications, space-filling curves have been used for pattern recognition [28], bandwidth reduction [29], image compression [1]–[3], and quantization of color spaces [4], [5]. We discuss the last two problems.

Images are two-dimensional arrays of data written in cells (or pixels) of a rectangular grid. To process images, one has to read the data in a sequence by scanning the image, visiting each cell of the grid. A scan can go simply over columns, rows,

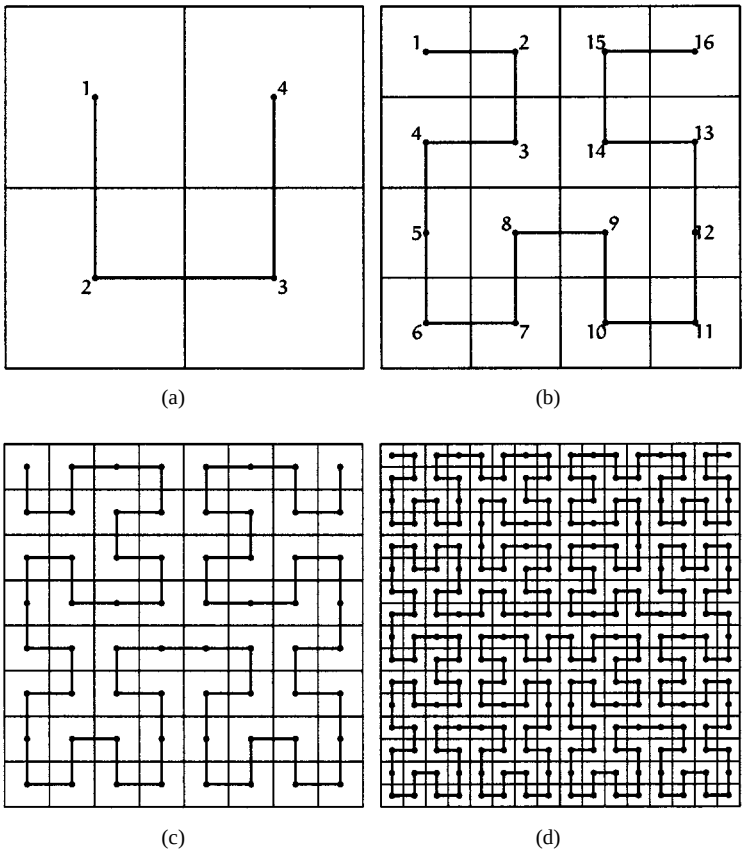


Fig. 10. Generating Hilbert's space-filling curve—the first four steps.

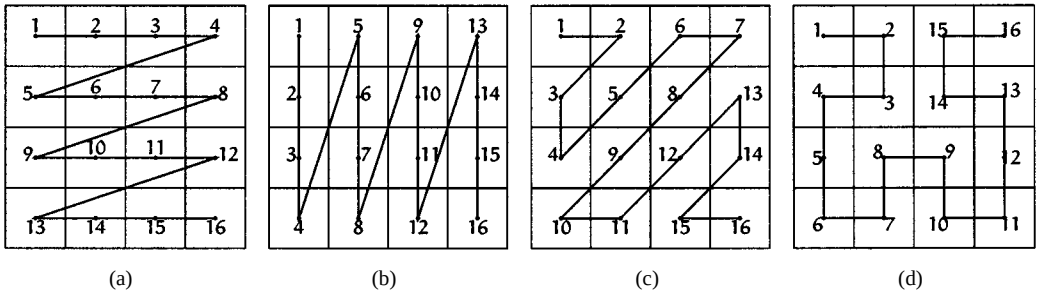


Fig. 11. Four ways to scan a two dimensional array of data.

or diagonals, as shown in Fig. 11(a)–(c) or as prescribed by a space-filling curve, such as the *Hilbert scan* in Fig. 11(d).

Information theoretically, images are two-dimensional arrays of random variables with finite spatial dependencies between them, which have to be accounted for in the one-dimensional scan output sequence. In order to be able to model the sequence as a random process with finite (short) memory, the scan visits to any two spatially close pixels should be close in time. In other words, numbers associated with pixels in a correlated neighborhood should not be far apart. Although no grid scan can have this property, the Hilbert scan in Fig. 11(d) visits every cell in the quadrant it enters before exiting the quadrant, and this remains true when the scan is extended to larger grids, as in Fig. 10. Note that this is a direct consequence of Hilbert's claim that when mapping an interval onto a square by partitioning into congruent replicas, the subsquares can be arranged so that adjacent subintervals correspond to adjacent subsquares with an edge in common.

In particular, the Hilbert scan covers arrays of size $2^k \times 2^k$, $k \geq 1$, quadrant by quadrant, never leaving a quadrant before visiting each of its cells, as illustrated in Fig. 10(d). Lempel and Ziv made use of this property to devise a way for lossless compression of images [1]. They considered compression schemes in which the image data is first read in a sequence by scanning the image, and then the resulting sequence is compressed by a finite-state encoder. It was shown that the best such scheme combines the Hilbert scanning path with the authors' compression algorithm for individual sequences presented in [27]. In other words, the compressibility of any picture is lower-bounded by that of the sequence derived from it under the Hilbert scan [1, Lemma 1]. A number of practical schemes for lossy image compression based upon space filling have been proposed in [2] and [3].

Digital algorithms for lossy compression of images are concerned not only with how a temporal sequence is formed from the two-dimensional image array, but also with the *alphabet* from which the sequence takes its values. The discrete alpha-



Fig. 12. A palette of 16 gray levels.

bets, commonly called palettes or color codebooks, are binary for black and white images and real-valued scalar for gray-scale images. A palette of 16 gray levels is shown in Fig. 12. For color images, palettes are obtained by sampling three-dimensional color spaces such as RGB or Lab. This process, referred to as *color quantization*, is extensively used to display, transfer, or store color images in Internet-based applications, computer graphics, and animation. Color quantization also has a big impact on processing of color-quantized images.

Techniques used to process image data such as edge detection, thresholding, or runlength encoding have been devised for sequences of scalars, and work well for gray-scale images. Color images are commonly represented as three sequences of real-valued scalars. Each sequence corresponds to one of the three components of the color space in use. Thus, for the RGB color space, the sequences correspond to the red, green, and blue color components. The components are usually processed independently until they are recombined for a color display. Color palettes where each color is described by a single number would make color images more amenable to the type of processing used for gray-scale ones. Attempts to construct such palettes (i.e., write a sequence in the space) have been made in [4], [5], and, more recently, in [7]. The first two proposals use space-filling curves. The third one uses spiral lattices, which are the subject of the next section.

IV. Two-Dimensional Cylindrical and Spiral Lattices

A. Phyllotaxy and Fibonacci Numbers

One of the ways that nature uses to write sequences on the plane is by cylindrical and spiral lattices. Fig. 13 illustrates a shell and how it might look if unwrapped, with points at centers of the white cells. Fig. 14 illustrates the seed-head of a daisy and how it might look if magnified, with numbers describing the *age* of the seed. The picture in Fig. 14(b) is made up of the first 90 points of what we shall refer to as the *Fibonacci spiral lattice*. The points are given by $z_n = \sqrt{n}\, e^{j2\pi \cdot n\tau}$, where $\tau = (\sqrt{5}+1)/2$ is the golden number and n is the number of time units corresponding to the age of the seed. Seeds begin their life at the origin 0.

B. Cylindrical Lattices in the Complex Plane

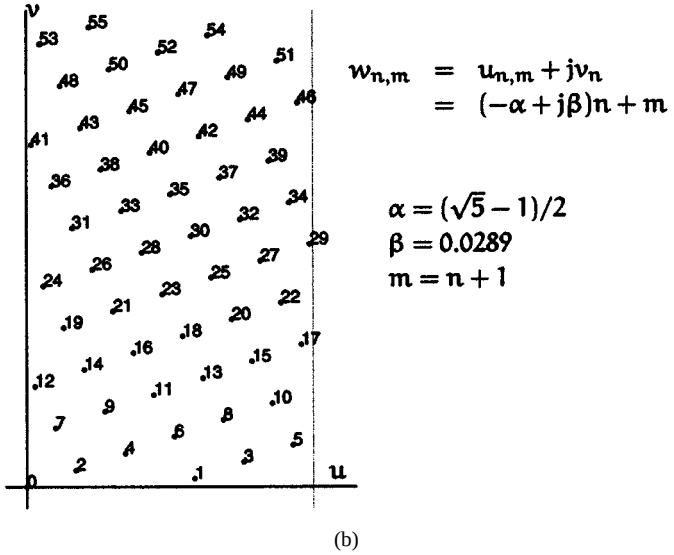
Sampling in two-dimensional Euclidean space is commonly represented by a *lattice*, see for example [31]. A two-dimensional lattice is specified by a pair $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ of linearly independent real vectors; the set of all points

$$\boldsymbol{x} = u_1\boldsymbol{a}_1 + u_2\boldsymbol{a}_2$$

with integers u_1 and u_2 comprise the lattice. A point on the lattice is specified by an ordered pair (u_1, u_2). As previously mentioned, the idea here is to devise a sampling scheme in which each point is described by a single number.



(a)



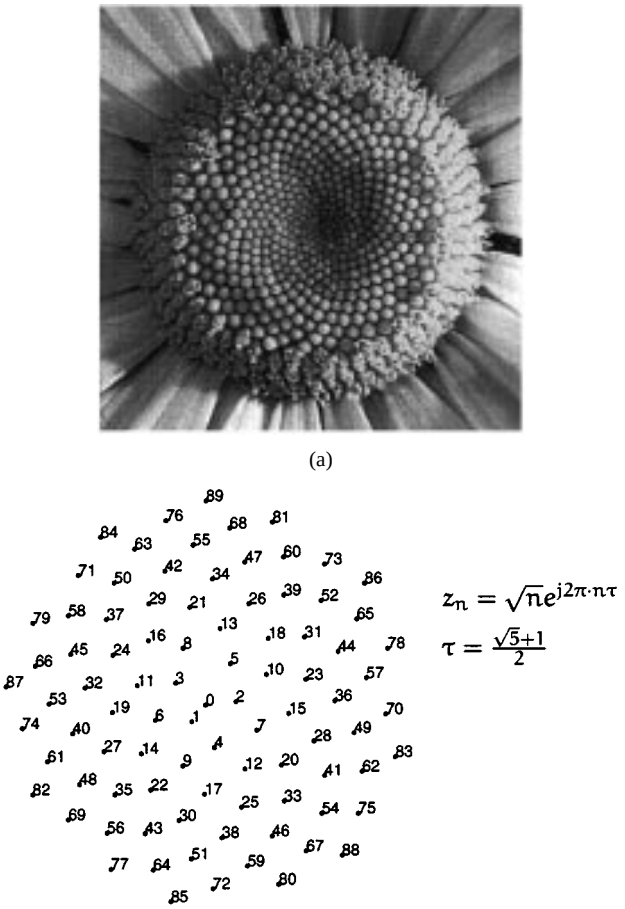
(b)

Fig. 13. A shell and a cylindrical lattice in the complex plain. (Photo: Scott Camazine, Pennsylvania State University [30].)

In two dimensions, we can equivalently consider the complex Euclidean plane $w = u + jv$. The points

$$w_{n,m} = u_{n,m} + jv_n = (-\alpha + j\beta)n + m, \qquad \alpha, \beta \in \boldsymbol{R},\ n, m \in \boldsymbol{Z}$$

form a lattice known as the *extended cylindrical lattice* [24]. Note that α and β are fixed constants defining the lattice, and again two integers, m and n, are necessary to specify a lattice point. Since we are concerned with sampling finite regions of the complex plane, we can (without loss of generality) consider only points $w_{n,m}$ for which $0 < u_{n,m} \le 1$, as shown in Fig. 13 for $\alpha = (\sqrt{5}-1)/2$ and $\beta = 0.0289$. It is easy to see that the requirement $0 < u_{n,m} \le 1$ translates into $m = \lceil n\alpha \rceil + 1$. Thus, each point can be identified only by n. Parameters α and β define the lattice, and thus the distribution of points in the plane. Two examples of cylindrical lattices are shown in Fig. 15. The



(a)

(b)

Fig. 14. The seed-head of a daisy and a spiral lattice in the complex plane. (Photo: Scott Camazine, Pennsylvania State University [30].)

"more irrational" the value of α the more uniform is the distribution of points. Precisely, the point distribution is determined by the continued fractions of α, which is discussed in Section IV-D.

C. Spiral Lattices in the Complex Plane

Because of the central symmetry of several popular color spaces, we consider the complex plane in polar coordinates and define a *spiral lattice* as a convenient means for sampling. A set of points

$$z_n = n^{\delta}\, e^{j2\pi \cdot n\tau}, \qquad \tau, \delta \in \boldsymbol{R},\ n \in \boldsymbol{Z}$$

constitute a spiral lattice. Fig. 14 shows a spiral lattice for $\tau = (\sqrt{5}+1)/2$ and $\delta = 1/2$, which we shall call *Fibonacci lattice*. Parameters τ and δ define the lattice, and thus the distribution of its points in the plane. For a given δ, parameter τ determines the axial distribution of the points. Examples of spiral lattices for two different values of τ and $\delta = 0.5$ are shown in Fig. 16. Note that for rational values of τ, e.g., $\tau = p/q$, the lattice points fall on rays radiating from the center at angles $l \cdot 2\pi p/q$, $0 \le l < q$. The axial point distribution is connected to the continued fractions of τ, discussed in Section IV-D. For a given τ, the parameter δ determines the radial distribution of the points. Examples of spiral lattices for two different values of δ and $\tau = (\sqrt{5}+1)/2$ are shown in Fig. 17.

D. Continued Fractions

W e shall call an expression of the form

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

a *continued fraction* with value x, and denote it by $[a_0; a_1, a_2, a_3, \ldots]$. The following theorem holds.

Theorem 1 [32, p. 16]: To every real number α, there corresponds a unique continued fraction with value equal to α. This fraction is finite if α is rational and infinite if α is irrational.

The rationals p_n/q_n given by

$$\frac{p_n}{q_n} = [a_0; a_1, a_2, a_3, \ldots a_n]$$

are known as *convergents*, or nth-step truncations, of the continued fraction expansion. The continued fraction expansion of the golden number is given by

$$\frac{\sqrt{5}+1}{2} = [1; 1, 1, \ldots]$$

and its convergents are ratios of successive Fibonacci numbers, i.e., $p_n/q_n = F_{n-1}/F_n$. *Noble numbers* are those numbers whose continued fraction expansion tails are all 1's, i.e., they agree with that of the golden ratio. The convergents p_n/q_n provide good rational approximations of α in the sense that there exists a constant c_α such that $|\alpha - p_n/q_n| < c_\alpha/q_n^2$. It turns out that $c_\alpha \le 1/\sqrt{5}$ for all α, with equality if and only if α is a noble number.

As discussed in Section IV-B, this rational approximation determines the axial distribution of the points $z_n = n^{\delta}\, e^{j2\pi \cdot n\alpha}$ of a spiral lattice. It is then not surprising that plants chose this (in a way) optimal packing. Although the hexagonal lattice provides the best sphere packing in two dimensions, because of the growth dynamics and circular shape of flower-heads, the seeds follow the spiral lattice rule. The set of bad irrationals (noble numbers) was extended by Markoff in 1879. He showed that all c_α's greater than $1/3$ take values in a discrete set of real numbers between $1/3$ and $1/\sqrt{5}$. This set of c_α's is called the *Markoff spectrum*, and the corresponding α's are called Markoff irrationals. For more on Markoff numbers and geometry, see for example [33]. Spiral lattices $z_n = \sqrt{n}\, e^{j2\pi \cdot n\alpha}$, where α is a Markoff irrational, can therefore be used for quantization of color spaces. For practical purposes of designing palettes with a finite number of colors, some rational numbers will do a good job as well.

E. Color Quantization and Processing by the Fibonacci Lattices

An entirely novel approach to sampling of the Lab color space, based on Fibonacci numbers and spiral lattices was proposed by Mojsilovic and Soljanin in [7]. The sampling scheme was used to generate color palettes for color quantization, display, and processing. Color image processing methods such as edge detection, differencing, averaging, and coding were described and implemented in [7]. Here, we highlight several important features of the color codebook structure, and then briefly explain how it can be used for processing of color images. We

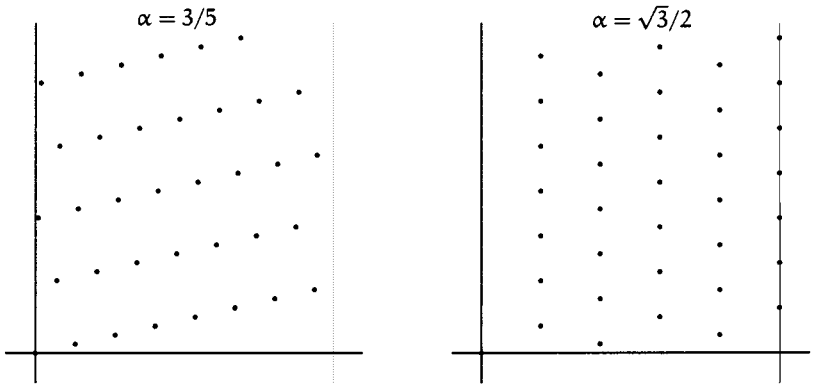


Fig. 15. Cylindrical lattices for two different values of α and $\beta = 0.03$.

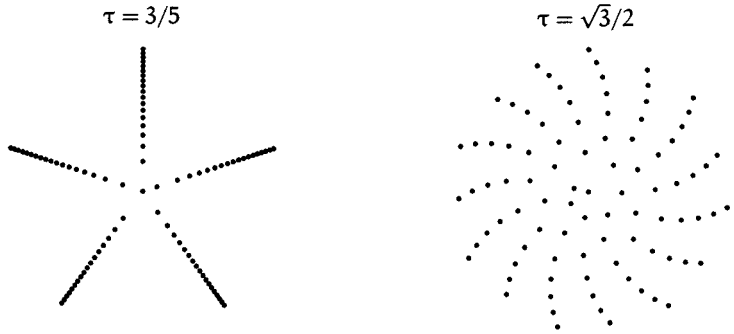


Fig. 16. Spiral lattices for two different values of τ and $\delta = 0.5$.

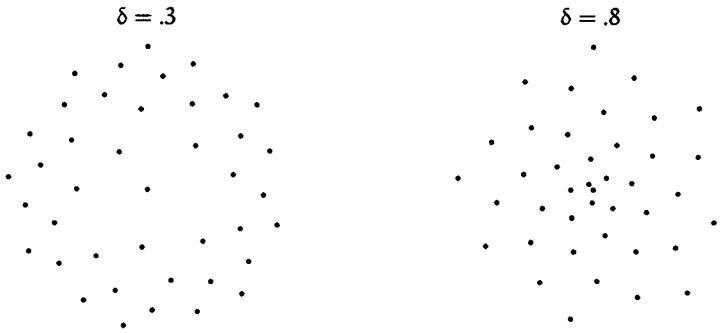
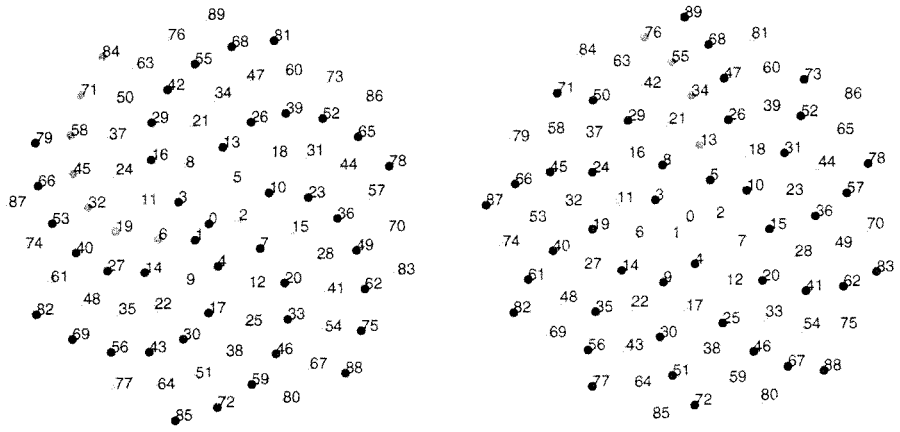


Fig. 17. Spiral lattices for two different values of δ and $\tau = (\sqrt{5}+1)/2$.

shall deal with the images given in the three-dimensional color space Lab, and explain only what is happening in a constant luminance cross section.

Let us look more closely at the sequence of points in Fig. 14(b) and in Fig. 18. The true source (growth pattern) of the sequence is a spiral winding around the center, connecting 0 with 1 and then 2, 3, 4, and so on up to 89. Observe the set of points again, forgetting for a moment about the main growth spiral. They appear to form two families of spirals sinking toward the center: one clockwise, one counterclockwise, as indicated in Fig. 18. It may be useful to look at the image of a daisy in Fig. 14(a), where the spiral patterns are more apparent. Notice that the numbers of neighboring points along clockwise spirals differ by 21, and along counterclockwise by 13. Now, choose a point and look at its neighbors in all directions. Note that the differences between the number of the chosen point and its neighbors are Fibonacci numbers. Pick another point and note that the differences between its number and its neighbors' numbers are again Fibonacci numbers, as for the first point.

This feature of the Fibonacci lattice allows simple and elegant solutions to several problems in color image processing [7]. Consider, for example, edge detection. Edge detection is a basic operation in image processing which examines how close are the colors of two adjacent pixels. In gray-scale images, this is naturally determined by the difference between the pixels' gray levels. Thus, if one of them has gray level 2 of the palette in



(a) (b)

Fig. 18. Spiral patterns in the Fibonacci lattice sinking counterclockwise and clockwise toward the center.

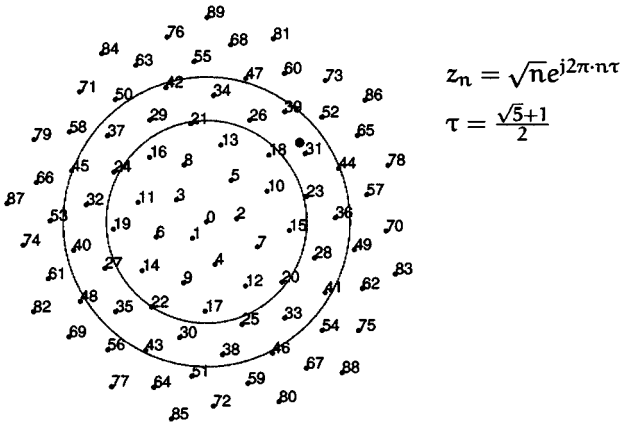


Fig. 19. Fast quantization.

Fig. 12 and the other gray level 10, then a difference of 8 and an edge are detected. In the Lab space, the Euclidean distance between two color points carries the information about their visual difference or similarity. But if the color image is quantized by a Fibonacci lattice, we can (as in gray-scale images) tell if there is an edge or not based only on the difference between the color numbers of two neighboring pixels; we just have to see whether the difference is a Fibonacci number or not.

The geometry and the numbering of the sample points in the color space significantly simplify otherwise computationally intensive operations of color quantization. Quantizers which minimize the *mean square error* (MSE) for an arbitrary point in the Lab space minimize the Euclidean distance between the point and the palette, i.e., search for the closest point in the palette. It is easy to see that computational requirements to perform this search within a codebook with little structure may be formidable. Since points of the Fibonacci lattice are given by $\sqrt{n}\, e^{j2\pi \cdot n\tau}$, a color whose squared distance to the origin is r^2 will be quantized to the palette point whose number n is close to r^2. The search for the closest point in the Fibonacci lattice is therefore limited to the a ring around the origin whose middle radius is approximately equal to r and whose width is slightly larger than the minimum distance between the lattice points, as shown in Fig. 19 where the larger point close to 31 is the one being quantized.

ACKNOWLEDGMENT

The author is grateful to the late A. D. Wyner who taught her what a great fortune it is to be able to do for a living what you would do for fun.

REFERENCES

[1] A. Lempel and J. Ziv, "Compression of two-dimensional data," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 2–8, Jan. 1986.
[2] A. Ansari and A. Fineberg, "Image data ordering and compression using Peano scan and lot," *IEEE Trans. Consumer Electron.*, vol. 38, pp. 436–445, June 1992.
[3] K.-M. Yang, L. Wu, and M. Mils, "Fractal based image coding scheme using Peano scan," in *Proc. 1988 IEEE Int. Symp. Circuits and Systems (ISCAS'88)*, Helsinki, Finland, June 1988, pp. 2301–2304.
[4] E. A. Patrick, D. R. Anderson, and F. K. Bechtel, "Mapping multidimensional space to one dimension for computer display," *IEEE Trans. Comput.*, vol. C-17, pp. 949–953, Oct. 1968.
[5] R. J. Stevens, A. F. Lehar, and F. H. Preston, "Manipulation and presentation of multidimensional image data using the Peano scan," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, pp. 520–526, Sept. 1983.
[6] E. Soljanin, "A Shannon-theoretic study of Penrose tilings," in *Proc. 1998 IEEE Int. Symp. Information Theory (ISIT'98)*, Cambridge, MA, Aug. 1998, p. 78.
[7] A. Mojsilovic and E. Soljanin, "Color quantization and processing by Fibonacci lattices," *IEEE Trans. Image Processing*, vol. 10, pp. 1712–1725, Nov. 2001.
[8] W. Weeks, IV and R. E. Blahut, "The capacity and coding gain of certain checkerboard codes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1193–1203, May 1998.
[9] P. H. Siegel and J. K. Wolf, "Bit-stuffing bounds on the capacity of 2-dimensional constrained arrays," in *Proc. 1998 IEEE Int. Symp. Information Theory (ISIT'98)*, Cambridge, MA, Aug. 1998, p. 323.
[10] K. A. Immink, P. H. Siegel, and J. K. Wolf, "Codes for digital recorders," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2260–2299, Oct. 1998.
[11] A. Kato and K. Zeger, "On the capacity of two-dimensional run-length constrained channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1527–1540, July 1999.
[12] ——, "Partial characterization of the positive capacity region of two-dimensional asymmetric run length constrained channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2666–2670, Nov. 2000.
[13] S. Forchhammer and J. Justesen, "Entropy bounds for constrained two-dimensional random fields," *IEEE Trans. Inform. Theory*, vol. 45, pp. 118–127, Jan. 1999.
[14] ——, "Bounds on the capacity of constrained two-dimensional codes," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2659–2666, Nov. 2000.
[15] N. J. Calkin and H. S. Wilf, "The number of independent sets in a grid graph," *SIAM J. Discr. Math.*, vol. 11, pp. 54–60, Feb. 1998.
[16] N. G. de Bruijn, "Algebraic theory of Penrose's nonperiodic tilings," *Nederl. Akad. Wetensch. Proc.*, ser. A 84, pp. 39–66, 1981.

[17] A. Connes, *Noncommutative Geometry*. San Diego, CA: Academic, 1994.
[18] R. M. Robinson, "Comments on the Penrose tiles," Mimeographed notes, Sept. 1975.
[19] B. Grünbaum and G. C. Shephard, *Tilings and Patterns*. New York: Freeman, 1987.
[20] M. Gardner, "Extraordinary nonperiodic tiling that enriches the theory of tiles," *Scient. Amer.*, pp. 110–121, Jan. 1977.
[21] M. Senechal, *Quasicrystals and Geometry*. New York: Cambridge Univ. Press, 1995.
[22] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamic and Coding*. New York: Cambridge Univ. Press, 1995.
[23] I. Stewart, *Life's Other Secret*. New York: Wiley, 1998.
[24] F. Rothen and A. Koch, "Phyllotaxis or the properties of spiral lattices," *J. Physique*, vol. 50, pp. 633–657, Mar. 1989.
[25] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 10, pp. 379–423, 623–656, Oct. 1948.
[26] H. Sagan, *Space-Filling Curves*. New York: Springer-Verlag, 1991.
[27] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 530–536, Sept. 1978.
[28] K. Abend, T. J. Harley, and L. N. Kanal, "Classification of binary random patterns," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 538–544, Oct. 1965.
[29] T. Bially, "Space-filling curves: Their generation and their application to bandwidth reduction," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 658–664, Nov. 1969.
[30] S. Camazine.. Penn State Univ., Dept. Entomology. [Online]. Available: http://www.psu.edu/dept/beehive
[31] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice Hall, 1995.
[32] A. Y. Khinchin, *Continued Fractions*. Mineola, NY: Dover, 1997.
[33] C. Series, "The geometry of Markoff numbers," *Math. Intell.*, vol. 7, no. 3, pp. 20–29, Mar. 1985.