

To appear in *Connection Science*
Vol. 00, No. 00, Month 20XX, 1–19

Analysis of Information Gain and Kolmogorov Complexity for Aesthetic Evaluation of Cellular Automata Configurations

Mohammad Ali Javaheri Javid^{a*} and Tim Blackwell^b and Robert Zimmer^c and Mohammad
Majid al-Rifaie^d

Department of Computing, Goldsmiths, University of London, UK ^{a,b,c,d}

(15 August 2015)

This paper addresses aesthetic problem in cellular automata, taking a quantitative approach for aesthetic evaluation. Although the Shannon's entropy is dominant in computational methods of aesthetics, it fails to discriminate accurately structurally different patterns in two-dimensions. We have adapted information gain measure and Kolmogorov complexity to overcome the shortcomings of entropic measures. The measures are customised to robustly quantify the complexity of multi-state cellular automata configurations.

Experiments are set up with different initial configurations in a two-dimensional multi-state cellular automaton whose corresponding structural and Kolmogorov complexity measures at global level are analysed. Preliminary outcomes on the resulting automata are promising, as they suggest the possibility of predicting the structural characteristics, symmetry and orientation of cellular automata generated patterns.

Keywords: Complexity, entropy, information gain, Kolmogorov complexity, computational aesthetics, cellular automata

1. Introduction

Cellular Automata (CA) were initially developed as a material independent framework to study the logic of self-reproducing behaviour of biological systems in the late 1940s by von Neumann and Ulam (Burks, 1970). In the 1960s the idea of using CA as artistic tool emerged from the works of Knowlton and Schwartz who produced “*Pixillation*”, one of the early computer generated animations (Knowlton, 1964; Schwartz & Schwartz, 1992). The “*Tapestry I*” and “*Tapestry II*”, still frames from “*Pixillation*” won the first prize of the *Eighth Annual Computer Art Contest* in 1970. The computer arts of Struycken (Struycken, 1976), Brown (Brown, 2001; Beddard & Dodds, 2009) and evolutionary architecture of Frazer (Frazer, 1995) are classical examples of CA based arts. Moreover, CA have been used for music composition, for example, Xenakis (Xenakis, 1992) and Miranda (Miranda, 2001).

The popularity of John Conway's Game of Life (Gardner, 1970) drew the attention of a wider community of researchers and digital artists to the unexplored potential of CA applications, especially in their capacity to generate complex behaviour from simple rules (Brown, July 1996). This fact has been noted by Wolfram, who himself produced some CA arts in the 1980s, “even a program that may have extremely simple rules will often be able to generate pictures that have striking aesthetic qualities-sometimes

*Corresponding author. Email: m.javaheri@gold.ac.uk

reminiscent of nature, but often unlike anything ever seen before” (Wolfram, 2002, p.11). He further emphasises that “one of the things I’ve been meaning to do is to make a bit more of a serious effort to use cellular automata in some kind of computer art” (Regis, 1988, p.250).

Although classical one-dimensional CA with binary states can exhibit complex behaviours, experiments with multi-state two-dimensional (2D) CA reveal a very rich spectrum of symmetric and asymmetric patterns which are extremely challenging to generate using conventional mathematical methods (Javaheri Javid & te Boekhorst, 2006; Javaheri Javid, al Rifaie, & Zimmer, 2014).

There are numerous studies on the quantitative (Langton, 1986) and qualitative behaviour (Wolfram, 1983, 1984, 2002) of CA but they are mostly concerned with categorising the rule space and the computational properties of CA. Since CA are one of the generative tools in computer art, means of evaluating the aesthetic qualities of CA generated patterns have a substantial contribution towards further automation of CA art. There have been some interesting attempts to develop means of controlling emergence of aesthetic behaviour in CA (W. Li, 1988, 1989; Sims, 1992; Mason, 1993; Ashlock & Tsang, 2009) but with less success. This is due to lack of computational methods in substituting human aesthetic perception speciality measures of complexity.

This work follows Birkhoff’s tradition in studying mathematical bases of aesthetics, especially the association of aesthetic judgement with the degree of order and complexity of a stimulus. Shannon’s information theory provided an objective measure of complexity. It led to emergence of various informational theories of aesthetics. However due to its nature, the entropic measures fail to take into account spacial characteristics of 2D patterns which is fundamental in addressing aesthetic problem in general and of CA generated configurations specifically.

In this paper, following our earlier works (Javaheri Javid et al., 2014; Javaheri Javid, al Rifaie, & Zimmer, 2015; Javaheri Javid, Blackwell, Zimmer, & al Rifaie, 2015), we investigate computational notion of aesthetic in CA as non-linear dynamical systems. Particularly, we examine *information gain* and *Kolmogorov complexity* as measures of complexity in multi-state 2D CA generated configurations.

This paper is organised as follows. Section 2 provides formal definitions and establishes notations of CA. Section 3 demonstrates that entropy is an inadequate measure of discriminating multi-state 2D CA configurations. In Section 4, in the framework of the objectives of this study a spatial complexity spectrum is formulated and the potential of information gain as a structural complexity measure is discussed. Section 5 provides formal notions of Kolmogorov complexity and method estimating it. Section 6 gives details of experiments that test the effectiveness of information gain and its relation to Kolmogorov complexity. The paper closes with a discussion and summary of findings.

2. Definition of Cellular Automata

DEFINITION 2.1 A cellular automaton is a regular tiling of a lattice with uniform deterministic finite state automata.

A cellular automaton \mathcal{A} is specified by a quadruple $\langle L, S, N, f \rangle$ where:

- L is a finite square lattice of cells (i, j) .
- $S = \{1, 2, \dots, k\}$ is set of states. Each cell (i, j) in L has a state $s \in S$.
- N is neighbourhood, as specified by a set of lattice vectors $\{e_a\}$, $a = 1, 2, \dots, N$. The neighbourhood of cell $r = (i, j)$ is $\{r + e_1, r + e_2, \dots, r + e_N\}$. A cell is considered

to be in its own neighbourhood so that one of $\{e_a\}$ is the zero vector $(0, 0)$. With an economy of notation, the cells in the neighbourhood of (i, j) can be numbered from 1 to N ; the neighbourhood states of (i, j) can therefore be denoted (s_1, s_2, \dots, s_N) . Periodic boundary conditions are applied at the edges of the lattice so that complete neighbourhoods exist for every cell in L .

- f is the update rule. f computes the state $s_1(t + 1)$ of a given cell from the states (s_1, s_2, \dots, s_N) of cells in its neighbourhood: $s_1(t + 1) = f(s_1, s_2, \dots, s_N)$. A quiescent state s_q satisfies $f(s_q, s_q, \dots, s_q) = s_q$.

Remark 1 There are two common neighbourhoods; a five-cell von Neumann neighbourhood $\{(0, 0), (\pm 1, 0), (0, \pm 1)\}$ and a nine-cell Moore neighbourhood $\{(0, 0), (\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1)\}$.

The collection of states for all cells in L is known as a *configuration* (C). The global rule F maps the whole automaton forward in time; it is the synchronous application of f to each cell. The behaviour of a particular \mathcal{A} is the sequence $c^0, c^1, c^2, \dots, c^{T-1}$, where c^0 is the initial configuration (IC) at $t = 0$.

CA behaviour is sensitive to the IC and to L, S, N and f . The behaviour is generally non-linear and sometimes very complex; no single mathematical analysis can describe, or even estimate, the behaviour of an arbitrary automaton. The vast size of the rule space, and the fact that this rule space is unstructured, mean that knowledge of the behaviour a particular cellular automaton, or even of a set of automata, gives no insight into the behaviour of any other CA. In the lack of any practical model to predict the behaviour of a cellular automaton, the only feasible method is to run simulations.

3. Aesthetics, Order and Complexity

Baumgarten (1714-1762) influenced by Descartes (1596-1650) in distinguishing between “*clear and distinct ideas*” from “*obscure and confused ideas*” argued that since thought and intellect types of cognitions are dealt with logic and reasoning, the perception and sensory types of cognitions should be dealt with a new science of sensory perception which he called *aesthetics*, deriving it from a Greek word for ‘things perceived by the senses’, as opposed to ‘things known by the mind’.

Aesthetic judgements have long been hypothesised to be correlated to the stimulus complexity and order (i.e. symmetry). Empirical studies on art and aesthetic preferences to address this problem initiated by Fechner who is also founded *experimental aesthetics* (Fechner, 1876). Wundt postulated that the stimulus complexity and appraisal are in an inverted-U curve relation which increases and then decreases with higher levels of arousal (Wundt, 1874). Birkhoff suggested an early mathematical model as *aesthetic measure* by arguing that the measure of aesthetic (M) (Eq. 1) is in direct relation with the degree of *order* (O) and in reverse relation with the *complexity* (C) of an object (Birkhoff, 1933),

$$M = \frac{O}{C}. \quad (1)$$

The validity of Birkhoff’s model, and his definition of order and complexity, has been challenged by empirical studies in (Wilson, July 1939). Eysenck conducted a series of experiments on Birkhoff’s model and suggested that a better expression of aesthetic evaluation function should consider a direct relation to stimulus complexity rather than

an inverse relation ($M = O \times C$) (Eysenck, 1941, 1942, 1968).

Information theory was developed in order to address communication over an unreliable channel (Shannon, 1948). Entropy is the core of this theory (Cover & Thomas, 2006). Let \mathcal{X} be discrete alphabet, X a discrete random variable, $x \in \mathcal{X}$ a particular value of X and $P(x)$ the probability of x . Then the entropy, $H(X)$, is:

$$H(X) = - \sum_{x \in \mathcal{X}} P(x) \log_2 P(x) \quad (2)$$

The quantity H is the average uncertainty in bits, $\log_2(\frac{1}{p})$ associated with X . Entropy can also be interpreted as the average amount of information needed to describe X . The value of entropy is always non-negative and reaches its maximum for the uniform distribution, $\log_2(|\mathcal{X}|)$:

$$0 \leq H \leq \log_2(|\mathcal{X}|). \quad (3)$$

The lower bound of relation (3) corresponds to a deterministic variable (no uncertainty) and the upper bound corresponds to a maximum uncertainty associated with a random variable. Entropy can be regarded as a measure of *order* and *complexity*. A low entropy implies low uncertainty and the message is highly predictable, ordered and less complex. High entropy implies a high uncertainty, less predictability, highly disordered and complex. These interpretations of entropy shaped the *informational theories of aesthetics*.

Moles (Moles, 1968), Bense (Bense, 1960; Bense & Nee, 1965; Bense, 1969) and Arnheim (Arnheim, 1954, 1966, 1969) were pioneers of the application of entropy to quantify order and complexity in Birkhoff's formula by adapting statistical measure of information in aesthetic objects. Bense argued that aesthetic objects are "vehicles of aesthetical information" where statistical information can quantify the aesthetical information of objects (Bense, 1960). His informational aesthetics has three basic assumptions. (1) Objects are material carriers of aesthetic state, and such aesthetic states are independent of subjective observers. (2) A particular kind of information is conveyed by the aesthetic state of the object (or process) as *aesthetic information* and (3) objective measure of aesthetic objects is in relation with degree of order and complexity in an object (Nake, 2012).

Herbert Franke put forward an *aesthetic perception* theory on the ground of *cybernetic aesthetics*. He made a distinction between the amount of information being stored and the rate of information flowing through a channel as *information flow* measured in *bits/sec* (Franke, 1977). His theory is based on psychological experiments which suggested that conscious working memory can not take more than 16 *bits/sec* of visual information. Then he argued that artists should provide a flow of information of about 16 *bits/sec* for works of art to be perceived as beautiful and harmonious. Staudek in his multi criteria approach (informational and structural) as *exact aesthetics* to Birkhoff's measure applied information flow I' by defining it as a measure assessing principal information transmission qualities in time. He used 16 *bits/sec* reference as channel capacity $C_r = 16 \text{ bits/sec}$ and a time reference of 8 seconds ($t_r = 8s$) to argue that artefacts with $I > 128 \text{ bits}$ will not fit into the conscious working memory for absorbing the whole aesthetic message (Staudek, 2002).

Berlyne, inspired by Wundt, suggested that the aesthetic judgement and pleasurable-ness of a stimulus starts at a relatively indifferent level, then increases as a function of complexity up to a certain level, then decreases and becomes more unpleasant as com-

plexity increases (Berlyne, 1960). Later experimental studies supported the aesthetic preference in relation to the complexity of visual and auditory stimuli conforming with an inverted-U curve (Munsinger & Kessen, 1964; Silvia, 2006). This principle explains the general dominance of stimuli or artworks that are not too simple and not too complex. Berlyne used an informational approach in his psychological experiments to determine human perceptual curiosity of visual figures (Berlyne, 1957).

Machado and Cardoso (Machado & Cardoso, 1998) proposed a model based on Birkhoff's approach as the ratio of *image complexity* to *processing complexity* by arguing that images with high visual complexity, are processed easily so they have the highest aesthetic value. Adapting Bense's informational aesthetics to different approaches of the concepts of order and complexity in an image in (Rigau, Feixas, & Sbert, 2007, 2008), three measures based on Kolmogorov complexity (M. Li, 1997), Shannon entropy (for RGB channels) and Zurek's physical entropy (Zurek, 1989) were introduced. Then the measures were applied to analyse aesthetic values of several paintings (Mondrian, Pollock, and van Gogh).

In summary:

- Birkhoff's aesthetic measure made *order* and *complexity* as integral parts of computational notions of aesthetics.
- Even though the validity of Birkhoff's approach in penalising complexity in aesthetic judgements has been contradicted by empirical studies, it seems that computational notions of aesthetics keep ignoring this well established fact.
- Despite the dominance of entropy as a measure of order and complexity, it fails to reflect on structural characteristics of 2D patterns. The main reason for this drawback is that it only reflects on the distribution of the symbols, and not on their spatial arrangements. This is in contrast to our intuitive perception of the complexity of patterns and is problematic for the purpose of measuring the complexity of 2D patterns. This fact also has been ignored on the developments of computational notions of aesthetics.

Fig. 1 illustrates the measure of entropy of 2D patterns with various structural characteristics with the uniform distribution of elements. Figs. 1a-b are patterns with ordered structures and Fig. 1c is a pattern with a fairly structureless random pattern.

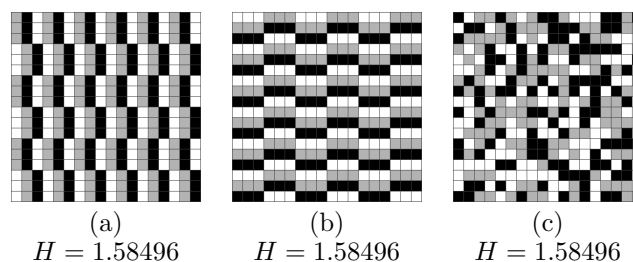


Figure 1. Measure of H for structurally different patterns with uniform distribution of elements.

The comparison of structural characteristics of these patterns with their corresponding entropy value shows that despite their structural differences, all of the patterns have the same entropy value. This clearly demonstrates the failure of entropy to discriminate structurally different 2D patterns. In other words entropy measure is invariant to spatial arrangement of composing elements of 2D patterns. This is in contrast to our intuitive perception of complexity of patterns. For the purpose of measuring complexity of CA behaviour specially with multi-state structures, it would be problematic if only entropy

measure is applied.

4. Information Gain Measure

Considering our intuitive perception of complexity and structural characteristics of 2D patterns, a complexity measure must be bounded by two extreme points of complete order and disorder. It is reasonable to assume that *regular structures*, *irregular structures* and *structureless* patterns lie along between these extremes, as illustrated in Fig. 2. A

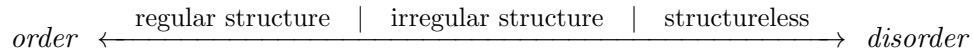


Figure 2. The spectrum of spatial complexity.

complete regular structure is a pattern of high symmetry, an irregular structure is a pattern with some sort of structure but not as regular as a fully symmetrical pattern and finally a structureless pattern is a random arrangement of elements (Javaheri Javid, Blackwell, et al., 2015). A measure introduced in (Bates & Shepard, 1993; Wackerbauer, Witt, Atmanspacher, Kurths, & Scheingraber, 1994; Andrienko, Yu. A., Brilliantov, N. V., & Kurths, J., 2000) and known as *information gain*, has been proposed as a means of characterising the complexity of dynamical systems and of 2D patterns. It measures the amount of information gained in bits when specifying the value, x , of a random variable X given knowledge of the value, y , of another random variable Y ,

$$G_{x,y} = -\log_2 P(x|y). \quad (4)$$

$P(x|y)$ is the conditional probability of a state x conditioned on the state y . Then the *mean information gain* (MIG), $\bar{G}_{X,Y}$, is the average amount of information gain from the description of the all possible states of Y :

$$\bar{G}_{X,Y} = \sum_{x,y} P(x,y) G_{x,y} = - \sum_{x,y} P(x,y) \log_2 P(x|y) \quad (5)$$

where $P(x,y)$ is the joint probability, $\text{prob}(X = x, Y = y)$. \bar{G} is also known as the conditional entropy, $H(X|Y)$ (Cover & Thomas, 2006). Conditional entropy is the reduction in uncertainty of the joint distribution of X and Y given knowledge of Y , $H(X|Y) = H(X,Y) - H(Y)$. The lower and upper bounds of $\bar{G}_{X,Y}$ are

$$0 \leq \bar{G}_{X,Y} \leq \log_2 |\mathcal{X}|. \quad (6)$$

DEFINITION 4.1 A structural complexity measure G , of a cellular automaton configuration is the sum of the mean information gains of cells having homogeneous/heterogeneous neighbouring cells over 2D lattice.

For a cellular automaton configuration, \bar{G} can be calculated by considering the distribution of cell states over pairs of cells r, s ,

$$\bar{G}_{r,s} = - \sum_{s_r, s_s} P(s_r, s_s) \log_2 P(s_r, s_s) \quad (7)$$

where s_r, s_s are the states at r and s . Since $|\mathcal{S}| = N$, $\bar{G}_{r,s}$ is a value in $[0, N]$.

The vertical, horizontal, primary diagonal (\backslash) and secondary diagonal ($/$) neighbouring pairs provide eight \overline{G} s; $\overline{G}_{(i,j),(i-1,j+1)}$, $\overline{G}_{(i,j),(i,j+1)}$, $\overline{G}_{(i,j),(i+1,j+1)}$, $\overline{G}_{(i,j),(i-1,j)}$, $\overline{G}_{(i,j),(i+1,j)}$, $\overline{G}_{(i,j),(i-1,j-1)}$, $\overline{G}_{(i,j),(i,j-1)}$ and $\overline{G}_{(i,j),(i+1,j-1)}$. The relative positions for non-edge cells are given by matrix M :

$$M = \begin{bmatrix} (i-1,j+1) & (i,j+1) & (i+1,j+1) \\ (i-1,j) & (i,j) & (i+1,j) \\ (i-1,j-1) & (i,j-1) & (i+1,j-1) \end{bmatrix}. \quad (8)$$

Correlations between cells on opposing lattice edges are not considered. The result of this edge condition is that $G_{i+1,j}$ is not necessarily equal to $\overline{G}_{i-1,j}$. In addition the differences between the horizontal (vertical) and two diagonal mean information rates reveal left/right (up/down), primary and secondary orientation of 2D patterns. So the sequence of generated configurations by a multi-state 2D cellular automaton can analysed by the differences between the vertical ($i, j \pm 1$), horizontal ($i \pm 1, j$), primary diagonal (P_d) and secondary diagonal (S_d) mean information gains by

$$\Delta \overline{G}_{i,j\pm 1} = |\overline{G}_{i,j+1} - \overline{G}_{i,j-1}| \quad (9a)$$

$$\Delta \overline{G}_{i\pm 1,j} = |\overline{G}_{i-1,j} - \overline{G}_{i+1,j}| \quad (9b)$$

$$\Delta \overline{G}_{P_d} = |\overline{G}_{i-1,j+1} - \overline{G}_{i+1,j-1}| \quad (9c)$$

$$\Delta \overline{G}_{S_d} = |\overline{G}_{i+1,j+1} - \overline{G}_{i-1,j-1}| \quad (9d)$$

Fig. 3 demonstrates the merits of \overline{G} in discriminating structurally different patterns for the sample patterns in Fig. 1. As it is evident, the measures of H are identical for structurally different patterns, however, the measure of \overline{G} s and $\Delta \overline{G}$ s are reflecting not only the complexity of patterns but their spatial arrangements too.

5. Kolmogorov Complexity of 2D Patterns

From information theory perspective, the object X has been a random variable drawn according to a probability mass function $P(x)$. If X is random, then the descriptive complexity of the event $X = x$ is $\log \frac{1}{P(x)}$, because $\lceil \log \frac{1}{P(x)} \rceil$ is the number of bits required to describe x . Thus the descriptive complexity of an object depends on the probability distribution (Cover & Thomas, 2006). Kolmogorov defined the algorithmic (descriptive) complexity of an object to the minimum length of a program such that a universal computer can generate a specific sequence (Kolmogorov, 1965). Thus, the Kolmogorov complexity of an object is independent of the probability distribution. Kolmogorov complexity is related to entropy ($H(X)$), in that the expected value of $K(x)$ for a random sequence is approximately the entropy of the source distribution for the process generating the sequence. However, Kolmogorov complexity differs from entropy in that it relates to the specific string being considered rather than the source distribution (M. Li, 1997;

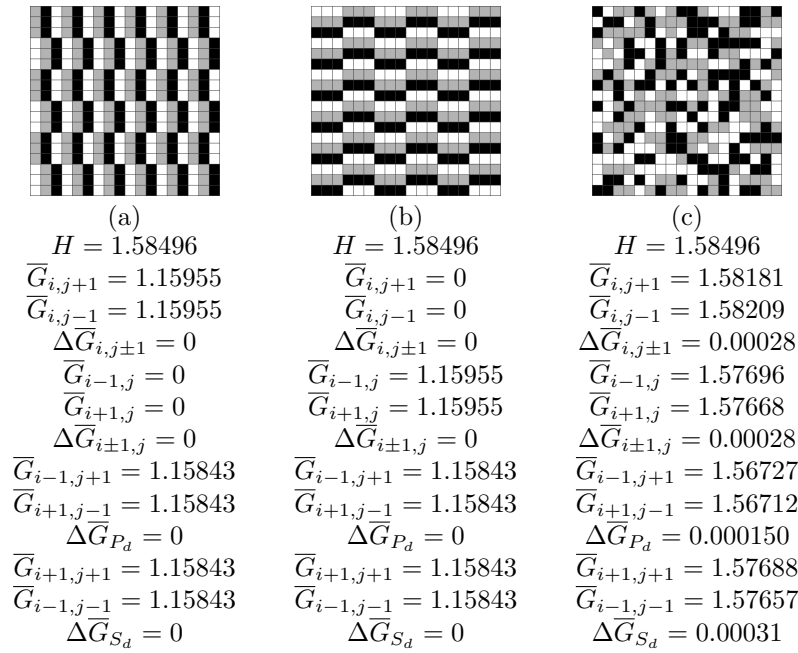


Figure 3. Measure of H , \bar{G} s and $\Delta\bar{G}$ s for structurally different patterns with uniform distribution of elements.

Cover & Thomas, 2006). Kolmogorov complexity can be described as follows, where φ represents a universal computer, p represents a program, and x represents a string:

$$K_\varphi(x) = \left\{ \min_{\varphi(p)=x} l(p) \right\} \quad (10)$$

Random strings have rather high Kolmogorov complexity on the order of their length, as patterns cannot be discerned to reduce the size of a program generating such a string. On the other hand, strings with a large amount of structure have fairly low complexity. Universal computers can be equated through programs of constant length, thus a mapping can be made between universal computers of different types. The Kolmogorov complexity of a given string on two computers differs by known or determinable constants. The Kolmogorov complexity $K(y|x)$ of a string y , given string x as input is described by the equation below

$$K_\varphi(y|x) = \left\{ \begin{array}{l} \min_{\varphi(p,y)=y} l(p) \\ \infty, \text{ if there is no } p \text{ such that } \varphi(p,x) = y \end{array} \right\} \quad (11)$$

where $l(p)$ represents program length p and φ is a particular universal computer under consideration. Thus, knowledge or input of a string x may reduce the complexity or program size necessary to produce a new string y . The major difficulty with Kolmogorov complexity is that you cannot compute it. Any program that produces a given string is an upper bound on the Kolmogorov complexity for this string, but you cannot compute the lower bound. Lempel and Ziv defined a measure of complexity for finite sequences rooted in the ability to produce strings from simple copy operations (Ziv & Lempel, 1978). This

method known as *LZ78* universal compression algorithm harnesses this principle to yield a universal compression algorithm that can approach the entropy of an infinite sequence produced by an ergodic source. So *LZ78* compression has been used as an estimator for K . Kolmogorov complexity is the ultimate compression bound for a given finite string, thus a natural choice for estimation of complexity is the class of universal compression techniques.

In order to estimate the K value of 2D configurations generated by multi-state CA, we generate linear strings of configurations by six different templates illustrates in Fig. 4.

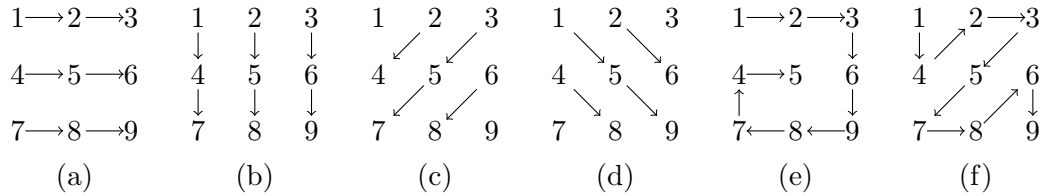


Figure 4. Six different templates.

Fig. 4a string horizontal $S_h = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, Fig. 4b vertical $S_v = \{1, 4, 7, 2, 5, 8, 3, 6, 9\}$, Fig. 4c diagonal $S_d = \{1, 2, 4, 3, 5, 7, 6, 8, 9\}$, Fig. 4d reverse diagonal $S_{rd} = \{3, 2, 6, 1, 5, 9, 4, 8, 7\}$, Fig. 4e spiral $S_s = \{1, 2, 3, 6, 9, 8, 7, 4, 5\}$, and Fig. 4f continuous spiral $S_{cs} = \{1, 4, 2, 3, 5, 7, 8, 6, 9\}$. Then using LZ78 compression algorithm, the upper bound of K is estimated. The comparison of the measurement of H , \overline{G}_s , $\Delta\overline{G}_s$ and K for structurally different patterns with uniform distribution of elements is illustrated in Fig. 5. As it is evident from the measurements, K is able to discriminate the complexity of patterns, however it fails to discriminate the spatial orientations.

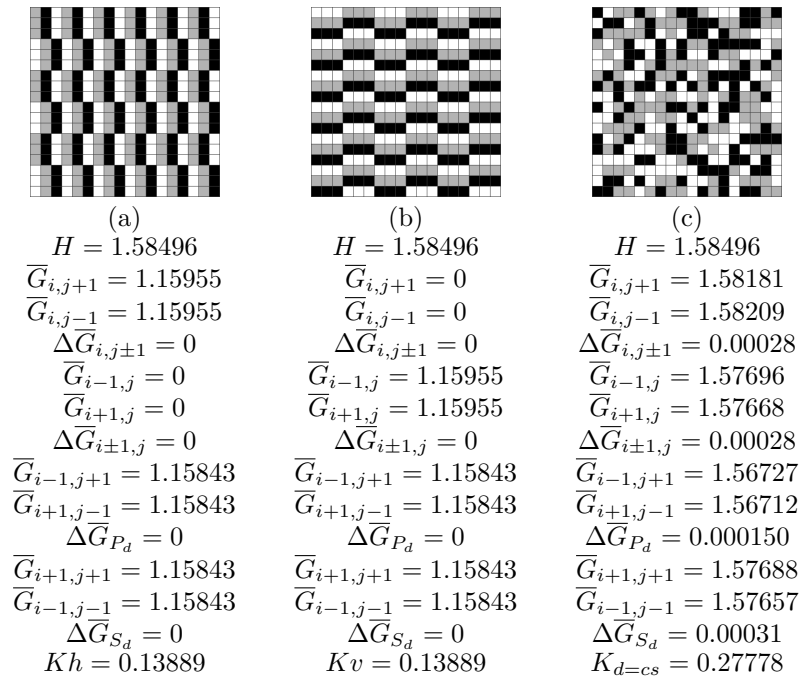


Figure 5. Measures of H , \overline{G}_s , $\Delta\overline{G}_s$ and K for structurally different patterns with uniform distribution of elements.

6. Experiments and Results

A set of experiments was designed to examine the effectiveness of \overline{G} in discriminating the particular configurations that are generated by a multi-state 2D cellular automaton. The experimental rule (Table 1) maps four states, represented by *white*, *red*, *blue* and *orange*; the quiescent state is *white*.

Table 1. Update rule of experimental cellular automaton.

$$\begin{aligned}
 L &= 65 \times 65 \text{ (4225 cells).} \\
 S &= \{0, 1, 2, 3\} \equiv \{\text{white}, \text{red}, \text{blue}, \text{orange}\} \\
 \mathcal{N} &: \text{Moore neighbourhood} \\
 f &: S^9 \mapsto S \\
 f(s_{i,j})(t) = s_{i,j}(t+1) &= \begin{cases} 1 & \text{if } s_{(i,j)}(t) = 0 & \text{and } \sigma = 1 \\ 3 & \text{if } s_{(i,j)}(t) = 1-3 & \text{and } \sigma = 2 \\ 2 & \text{if } s_{(i,j)}(t) = 1-3 & \text{and } \sigma = 3 \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

where σ is the sum total of the neighbourhood states.

The experiments are conducted with four different ICs: (1) all quiescent states cells except for a single cell (Fig. 6a) (2) a right oriented 5 cell (Fig. 6b), (3) a left oriented 5 cell (Fig. 6c) and (4) a random configuration with 2112 *white* quiescent states cells covering $\approx 50\%$ of the lattice, 749 *red*, 682 *blue* and 682 *orange* cells (Fig. 6d). The experimental rule has been iterated synchronously for 150 successive time steps. Figs. 7, 9, 11 and 13 illustrate a sample of time steps starting from four different ICs. Then the sequence of configurations are analysed by 9a, 9b, 9c, 9d and K .

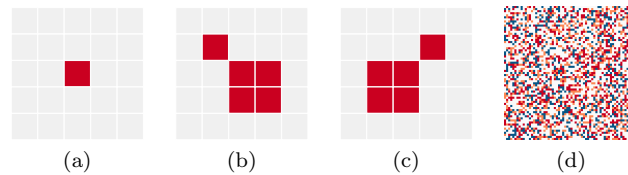


Figure 6. The four different ICs.

The behaviour of cellular automaton from the single cell IC is a sequence of symmetrical patterns (Fig. 7). This fact has been reflected on the measurements of $\Delta\overline{G}s$ (Fig. 8), where they are constant for the 150 time steps ($\Delta\overline{G}_{i,j\pm 1} = \Delta\overline{G}_{i\pm 1,j} = \Delta\overline{G}_{P_d} = \Delta\overline{G}_{S_d} = 0$). This is an indicator of the development of complete symmetrical patterns in four directions for each of 150 configurations generated by experimental cellular automaton. However, the measurement of entropy starts from $H_0 = 0.00319$ and reaches $H_{150} = 1.47979$ at the end of the runs.

The behaviour of cellular automaton from two 5 cell ICs (6b and 6c) are sequence of symmetrical patterns with different orientations (Fig. 14). The measurements of H for these two sequences of structurally different but symmetrical configurations are identical from $t = 0$ to $t = 150$, where $H_0^{6b} = H_0^{6c} = 0.01321$ and $H_{150}^{6b} = H_{150}^{6c} = 1.43241$ (Fig. 16). On the other hand the measurements of $\Delta\overline{G}s$ especially $\Delta\overline{G}_{P_d}$ and $\Delta\overline{G}_{S_d}$ are reflecting the differences in the orientations of symmetrical configurations (Figs. 14). This is further illustrated in Fig. 15 where the measures of H , $\overline{G}s$ and $\Delta\overline{G}s$ are compared for two configurations generated at $t = 40$ from two different 6b and 6c ICs.

[a side by side comparison of the measurements of IG for and H]

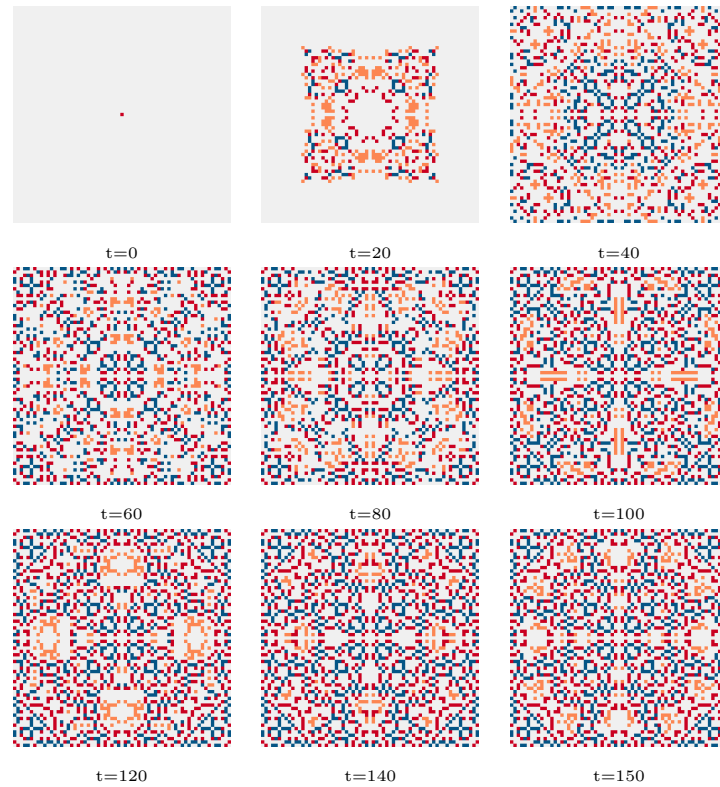


Figure 7. Space-time diagram of the experimental cellular automaton for sample time steps starting from the single cell IC (6a).

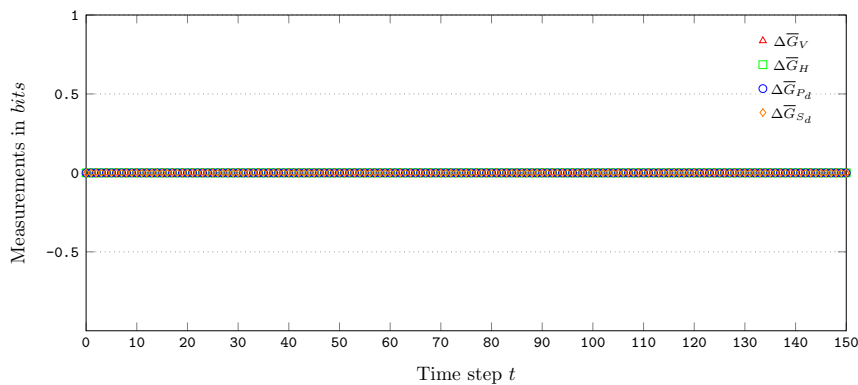


Figure 8. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 6a ICs.

The behaviour of cellular automaton from the random IC is a sequence of irregular structures (Fig. 13). The formation of patterns with local structures has reduced the values of $\Delta\overline{G}s$ until a stable oscillating pattern is attained. This is an indicator of the development of irregular structures. However the patterns are not random patterns since the maximum four-state value $\log_2(4) = 2$ (Eq. 6).

These experiments demonstrate that a cellular automaton rule seeded with different ICs leads to the formation of patterns with structurally diverse characteristics. The gradient of the mean information rate along lattice axes is able to detect the structural characteristics of patterns generated by this particular multi-state 2D cellular automaton. From the comparison of H with $\Delta\overline{G}s$ in the set of experiments, it is clear that entropy fails

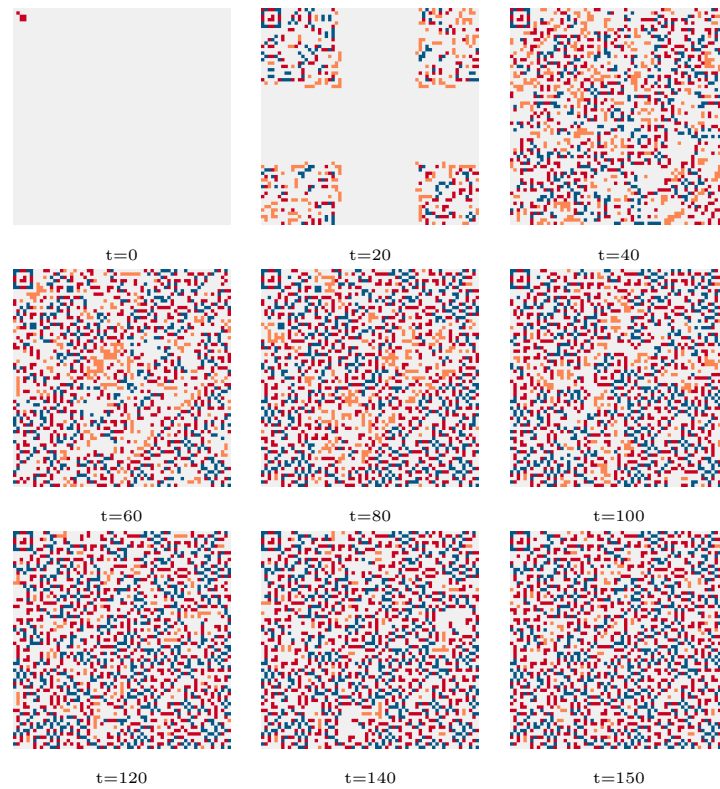


Figure 9. Space-time diagram of the experimental cellular automaton for sample time steps starting from the 6b IC.

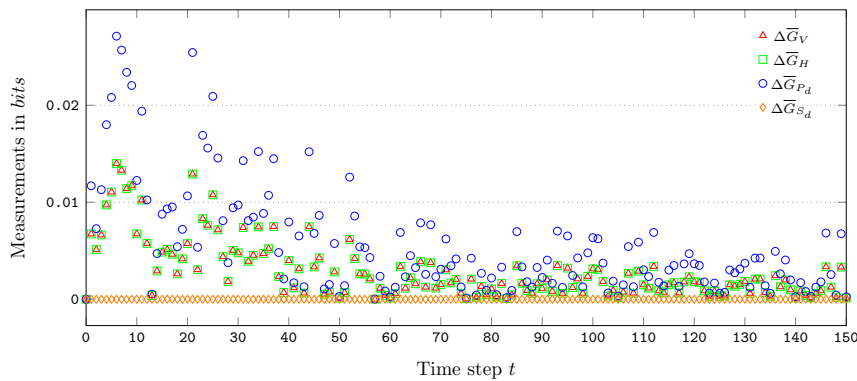


Figure 10. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 6b and 6c ICs.

to discriminate between the diversity of patterns that can be generated by various CA. The structured but asymmetrical patterns emerging from the random start are clearly distinguished from the symmetrical patterns including their orientation. As it is evident from the results of experiments, the measures of H are identical for structurally different patterns, however, the measure of $\overline{G}s$ and $\Delta\overline{G}s$ are reflecting not only the complexity of patterns but their spatial arrangements (i.e. orientation of symmetries) as well.

The relationship between K and $\overline{G}s$ are examined by Eq. 12 (Pearson correlation coefficient). Table 2 illustrates the calculations of r for different directional $\overline{G}s$. Since the r are ≈ 0.99 , so there are strong positive correlation between K and $\overline{G}s$.

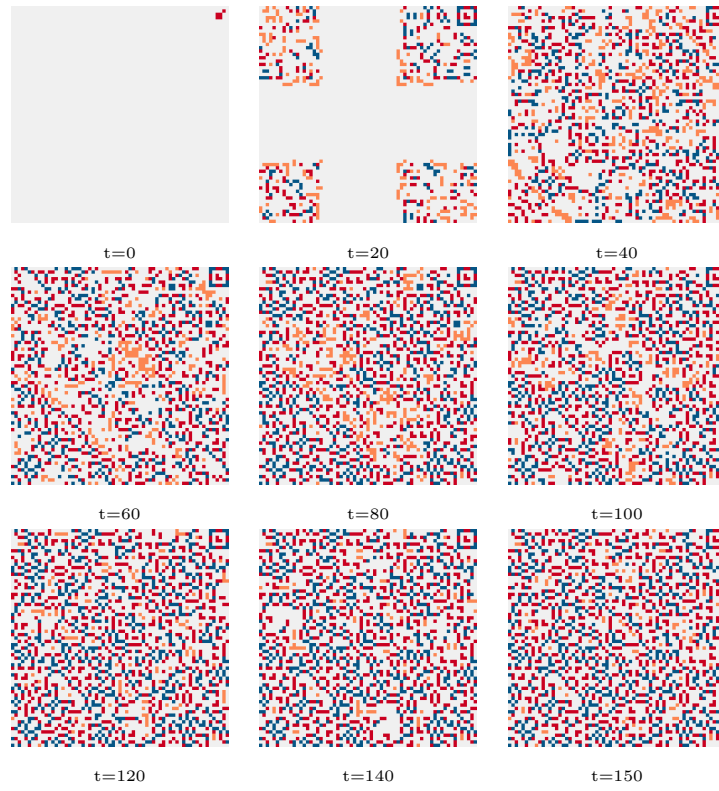


Figure 11. Space-time diagram of the experimental cellular automaton for sample time steps starting from the 6c IC.

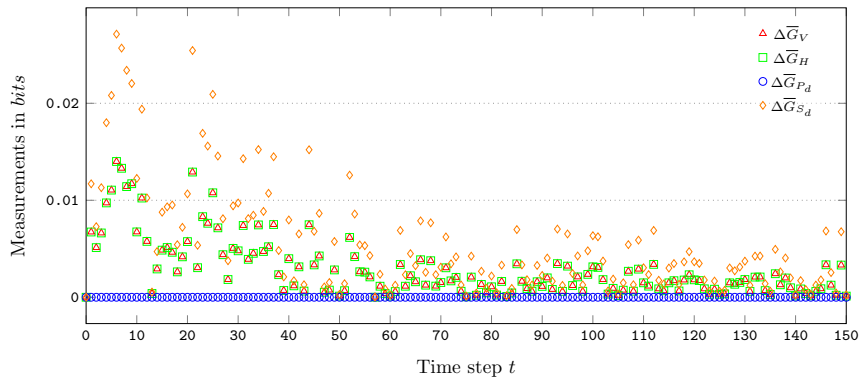


Figure 12. The measurements of $\Delta \bar{G}$ s for 150 time steps starting from 6b and 6c ICs.

$$r = r_{KG} = \frac{\sum_{i=1}^{150} (K_i - \bar{K})(G_i - \bar{G})}{\sqrt{\sum_{i=1}^{150} (K_i - \bar{K})^2} \sqrt{\sum_{i=1}^{150} (G_i - \bar{G})^2}} \quad (12)$$

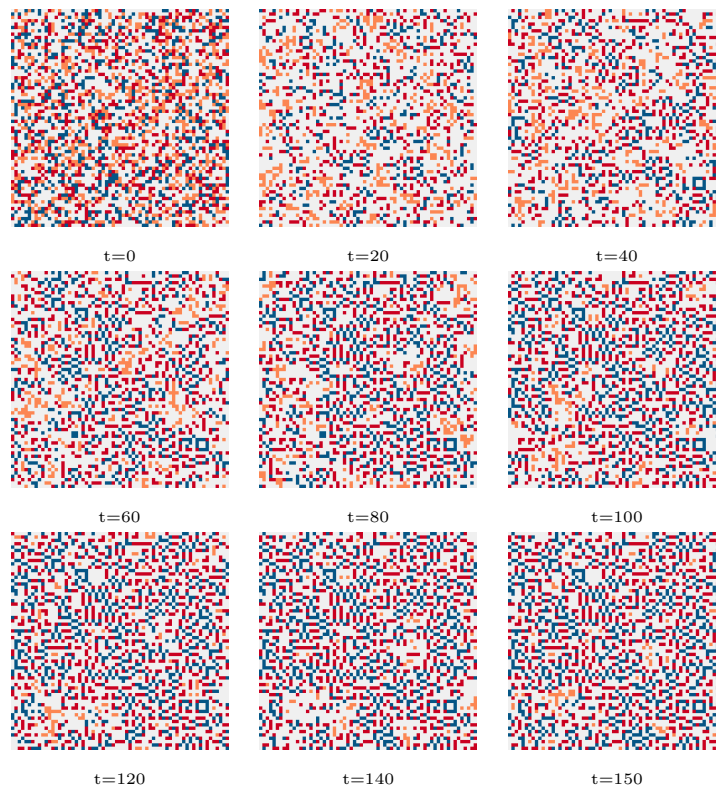


Figure 13. Space-time diagram of the experimental cellular automaton for sample time steps starting from the random (6d) IC.

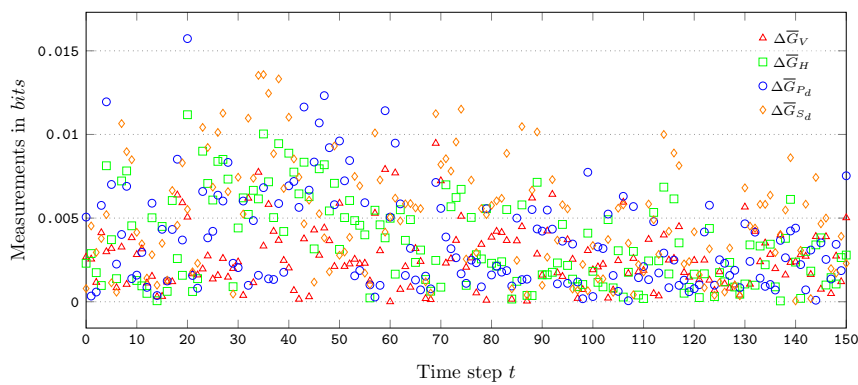
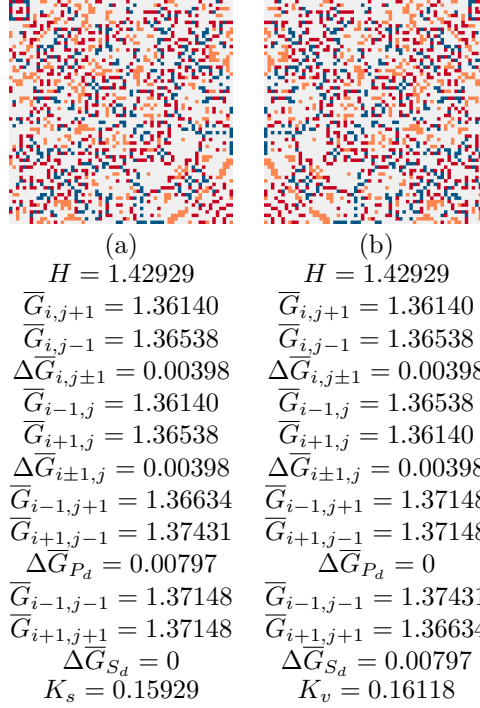
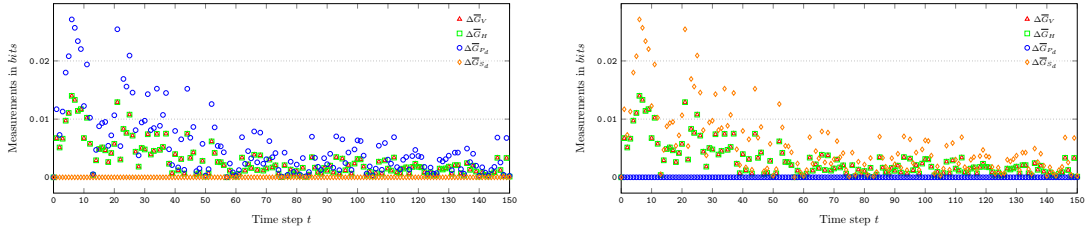


Figure 14. The measurements of $\Delta\bar{G}$ s for 150 time steps starting from 6b and 6c ICs.

Figure 15. The comparison of H , $\bar{G}s$ and $\Delta\bar{G}s$ at $t = 40$ for 6b (a) and 6c (b) ICs.Figure 16. The measurements of $\Delta\bar{G}s$ for 150 time steps starting from 6b and 6c ICs.Table 2. Calculations of r for different ICs.

$r_{K\bar{G}_{i,j+1}} = 0.9985$	$r_{K\bar{G}_{i,j-1}} = 0.9985$	$r_{K\bar{G}_{i-1,j}} = 0.9985$	$r_{K\bar{G}_{i+1,j}} = 0.9985$
$r_{K\bar{G}_{i-1,j+1}} = 0.9975$	$r_{K\bar{G}_{i+1,j-1}} = 0.9975$	$r_{K\bar{G}_{i+1,j+1}} = 0.9975$	$r_{K\bar{G}_{i-1,j-1}} = 0.9975$

Calculations of r for 6a IC.

$r_{K\bar{G}_{i,j+1}} = 0.9996$	$r_{K\bar{G}_{i,j-1}} = 0.9995$	$r_{K\bar{G}_{i-1,j}} = 0.9996$	$r_{K\bar{G}_{i+1,j}} = 0.9995$
$r_{K\bar{G}_{i-1,j+1}} = 0.9996$	$r_{K\bar{G}_{i+1,j-1}} = 0.9996$	$r_{K\bar{G}_{i+1,j+1}} = 0.9994$	$r_{K\bar{G}_{i-1,j-1}} = 0.9995$

Calculations of r for 6b IC.

$r_{K\bar{G}_{i,j+1}} = 0.9996$	$r_{K\bar{G}_{i,j-1}} = 0.9995$	$r_{K\bar{G}_{i-1,j}} = 0.9995$	$r_{K\bar{G}_{i+1,j}} = 0.9996$
$r_{K\bar{G}_{i-1,j+1}} = 0.9995$	$r_{K\bar{G}_{i+1,j-1}} = 0.9996$	$r_{K\bar{G}_{i+1,j+1}} = 0.9996$	$r_{K\bar{G}_{i-1,j-1}} = 0.9996$

Calculations of r for 6c IC.

$r_{K\bar{G}_{i,j+1}} = 0.9854$	$r_{K\bar{G}_{i,j-1}} = 0.9842$	$r_{K\bar{G}_{i-1,j}} = 0.9874$	$r_{K\bar{G}_{i+1,j}} = 0.9838$
$r_{K\bar{G}_{i-1,j+1}} = 0.9885$	$r_{K\bar{G}_{i+1,j-1}} = 0.9879$	$r_{K\bar{G}_{i+1,j+1}} = 0.9794$	$r_{K\bar{G}_{i-1,j-1}} = 0.9831$

Calculations of r for 6d IC.

7. Conclusion

CA are powerful tools for the pattern generation and have been used as a generative tool in computer art. Indeed, multi-state 2D CA can generate aesthetically appealing and complex patterns with various structural characteristics. Entropy, which is a statistical measure of the distribution of cell states, is not in general able to distinguish structurally different configurations generated by CA. However mean information gain, takes into account conditional and joint probabilities between pairs of cells and, since it is based on correlations between cells, holds promise for patterns discrimination. Kolmogorov algorithmic complexity is another measure of complexity which is used to estimate the complexity of 2D configurations generated by a cellular automaton.

This paper reports on a set of experiments for four different initial conditions of a cellular automaton. The potential of mean information gain and Kolmogorov complexity for distinguishing multi-state 2D CA patterns is demonstrated. So the measure appears to be particularly good at distinguishing different kinds of random patterns from non-random patterns however, information gain can also discriminate the orientation of symmetries as well. Since CA are one of the generative tools in computer art, means of evaluating the aesthetic qualities of CA generated patterns could have a substantial contribution towards further automation of aesthetic behaviour generation-evaluation of CA art.

References

- Andrienko, Yu. A., Brilliantov, N. V., & Kurths, J. (2000). Complexity of two-dimensional patterns. *Eur. Phys. J. B*, 15(3), 539–546.
- Arnheim, R. (1954). *Art and visual perception: A psychology of the creative eye*. Univ of California Press.
- Arnheim, R. (1966). Towards a psychology of art/entropy and art an essay on disorder and order. *The Regents of the University of California*.
- Arnheim, R. (1969). *Visual thinking*. Univ of California Press.
- Ashlock, D., & Tsang, J. (2009). Evolved art via control of cellular automata. In *Evolutionary computation, 2009. cec'09. IEEE congress on* (pp. 3338–3344).
- Bates, J. E., & Shepard, H. K. (1993). Measuring complexity using information fluctuation. *Physics Letters A*, 172(6), 416–425.
- Beddard, H., & Dodds, D. (2009). *Digital Pioneers*. V&A Publishing.
- Bense, M. (1960). *Aesthetica: Programmierung des schönen, allgemeine texttheorie und textästhetik [aesthetica : Programming of beauty, general text theory and aesthetics]*. Agis-Verlag.
- Bense, M. (1969). *Kleine abstrakte ästhetik [small abstract aesthetics]* (Vol. 38). E. Walther.
- Bense, M., & Nee, G. (1965). Computer grafik. In M. Bense & E. Walther (Eds.), *Edition rot* (Vol. 19). Walther, Stuttgart.
- Berlyne, D. E. (1957). Conflict and information-theory variables as determinants of human perceptual curiosity. *Journal of experimental psychology*, 53(6), 399.
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. McGraw-Hill Book Company.
- Birkhoff, G. (1933). *Aesthetic measure*. Harvard University Press.
- Brown, P. (2001). Stepping stones in the mist. In *Creative evolutionary systems* (pp. 387–407).
- Brown, P. (July 1996). Emergent behaviours: Towards computational aesthetics. *Artlink*, 16(2 & 3).
- Burks, A. W. (1970). *Essays on cellular automata*. University of Illinois Press.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory (wiley series in telecommunications and signal processing)*. Wiley-Interscience.
- Eysenck, H. J. (1941). The empirical determination of an aesthetic formula. *Psychological Review*, 48(1), 83.

- Eysenck, H. J. (1942). The experimental study of the ‘good gestalt’ – a new approach. *Psychological Review*, 49(4), 344.
- Eysenck, H. J. (1968). An experimental study of aesthetic preference for polygonal figures. *The Journal of General Psychology*, 79(1), 3–17.
- Fechner, G. (1876). *Vorschule der aesthetik [preschool of aesthetics]* (Vol. 1-2). Breitkopf & Härtel.
- Franke, H. W. (1977). A cybernetic approach to aesthetics. *Leonardo*, 10(3), 203–206.
- Frazer, J. (1995). *An evolutionary architecture*. Architectural Association Publications, Themes VII.
- Gardner, M. (1970, October). Mathematical games - the fantastic combinations of john conway’s new solitaire game, life. *Scientific American*, 120–123.
- Javaheri Javid, M. A., al Rifaie, M. M., & Zimmer, R. (2014). Detecting Symmetry in Cellular Automata Generated Patterns Using Swarm Intelligence. In A.-H. Dediu, M. Lozano, & C. Martín-Vide (Eds.), *Theory and Practice of Natural Computing* (Vol. 8890, pp. 83–94). Springer International Publishing.
- Javaheri Javid, M. A., al Rifaie, M. M., & Zimmer, R. (2015). An Informational Model for Cellular Automata Aesthetic Measure. In *Aisb 2015 symposium on computational creativity*. University of Kent, Canterbury, UK.
- Javaheri Javid, M. A., Blackwell, T., Zimmer, R., & al Rifaie, M. M. (2015). Spatial Complexity Measure for Characterising Cellular Automata Generated 2D Patterns. In Francisco Pereira, Penousal Machado, Ernesto Costa and Amílcar Cardoso (Ed.), *Progress in artificial intelligence, 17th portuguese conference on artificial intelligence, epia 2015, coimbra, portugal, september 8-11, 2015, proceedings.* (Vol. 9273). Springer International Publishing.
- Javaheri Javid, M. A., & te Boekhorst, R. (2006). Cell Dormancy in Cellular Automata. In V. N. Alexandrov, G. D. van Albada, P. M. A. Sloom, & J. Dongarra (Eds.), *International Conference on Computational Science (3)* (Vol. 3993, pp. 367–374). Springer.
- Knowlton, K. C. (1964). A computer technique for producing animated movies. In *Proceedings of the april 21-23, 1964, spring joint computer conference* (pp. 67–87).
- Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of information transmission*, 1(1), 1–7.
- Langton, C. G. (1986). Studying artificial life with cellular automata. *Physica D: Nonlinear Phenomena*, 22(1), 120–149.
- Li, M. (1997). *An introduction to kolmogorov complexity and its applications*. Springer.
- Li, W. (1988). *Pretty pictures generated by two-state five-neighbor cellular automata* (Tech. Rep.). CCSR Technical Report No. 15., University of Illinois.
- Li, W. (1989). Complex patterns generated by next nearest neighbors cellular automata. *Computers & Graphics*, 13(4), 531–537.
- Machado, P., & Cardoso, A. (1998). Computing aesthetics. In *Advances in artificial intelligence* (pp. 219–228). Springer.
- Mason, W. K. (1993). Art from cellular automata and symmetrized dot-patterns. *Computers & graphics*, 16(4), 439–441.
- Miranda, E. (2001). *Composing music with computers* (No. v. 1). Focal Press.
- Moles, A. (1968). *Information theory and esthetic perception. trans. je cohen*. U. Illinois Press.
- Munsinger, H., & Kessen, W. (1964). Uncertainty, structure, and preference. *Psychological Monographs: General and Applied*, 78(9), 1.
- Nake, F. (2012). Information aesthetics: An heroic experiment. *Journal of Mathematics and the Arts*, 6(2-3), 65–75.
- Regis, E. (1988). *Who got einstein’s office?: eccentricity and genius at the institute for advanced study*. Basic Books.
- Rigau, J., Feixas, M., & Sbert, M. (2007). Conceptualizing birkhoff’s aesthetic measure using shannon entropy and kolmogorov complexity. In D. W. Cunningham, G. Meyer, & L. Neumann (Eds.), *Workshop on computational aesthetics* (pp. 105–112). Banff, Alberta, Canada: Eurographics Association.

- Rigau, J., Feixas, M., & Sbert, M. (2008). Informational aesthetics measures. *Computer Graphics and Applications, IEEE*, 28(2), 24–34.
- Schwartz, L., & Schwartz, L. (1992). *The computer artist's handbook: Concepts, techniques, and applications*. W W Norton & Company Incorporated.
- Shannon, C. (1948, October). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27, 379–423 & 623–656.
- Silvia, P. J. (2006). *Exploring the psychology of interest*. Oxford University Press.
- Sims, K. (1992). Interactive evolution of dynamical systems. In *Toward a practice of autonomous systems: Proceedings of the first european conference on artificial life* (pp. 171–178).
- Stauderk, T. (2002). *Exact Aesthetics. Object and Scene to Message* (Unpublished doctoral dissertation). Faculty of Informatics, Masaryk University of Brno.
- Struycken, P. (1976). Splash 1972/1974. In R. Leavitt (Ed.), *Artist and computer* (pp. 30–31). Harmony Books.
- Wackerbauer, R., Witt, A., Atmanspacher, H., Kurths, J., & Scheingraber, H. (1994). A comparative classification of complexity measures. *Chaos, Solitons & Fractals*, 4(1), 133–173.
- Wilson, D. J. (July 1939). An experimental investigation of Birkhoff's aesthetic measure. *The Journal of Abnormal and Social Psychology*, 34(3), 390.
- Wolfram, S. (1983). Statistical mechanics of cellular automata. *Reviews of modern physics*, 55(3), 601–644.
- Wolfram, S. (1984). Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, 10(1), 1–35.
- Wolfram, S. (2002). *A new kind of science*. Wolfram Media Inc.
- Wundt, W. M. (1874). *Grundzüge de physiologischen psychologie [outline of physiological psychology]*. W. Engelman, Leipzig, Germany.
- Xenakis, I. (1992). *Formalized music: thought and mathematics in composition*. Pendragon Press.
- Ziv, J., & Lempel, A. (1978). Compression of individual sequences via variable-rate coding. *Information Theory, IEEE Transactions on*, 24(5), 530–536.
- Zurek, W. H. (1989). Algorithmic randomness and physical entropy. *Physical Review A*, 40(8), 4731.