Information Gain Measure for Structural Discrimination of Cellular Automata Configurations

Mohammad Ali Javaheri Javid*, Tim Blackwell†, Robert Zimmer‡, Mohammad Majid al-Rifaie§

Department of Computing

Goldsmiths, University of London

London SE14 6NW, UK

*m.javaheri, †t.blackwell, ‡r.zimmer, §m.majid@gold.ac.uk

Abstract—Cellular automata (CA) are known for their capability in exhibiting interesting emergent behaviour and capacity to generate complex and often aesthetically appealing patterns through the local interaction of rules. Mean information gain has been suggested as a measure of discriminating structurally different two-dimensional (2D) patterns. This paper addresses quantitative evaluation of the complexity of CA generated configurations. In particular, we examine information gain as a spatial complexity measure for discriminating multi-state 2D CA generated configurations. This information-theoretic quantity, also known as conditional entropy, takes into account conditional and joint probabilities of cell states in a 2D plane. The effectiveness of the measure is shown in a series of experiments for multi-state 2D patterns generated by CA. The results of the experiments show that the measure is capable of distinguishing the structural characteristics including symmetries and randomness of 2D CA patterns.

I. INTRODUCTION

Cellular Automata (CA) were developed in the late 1940s by von Neumann and Ulam as a material independent framework to simulate the self-reproducing behaviour of biological systems. 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 [1], [2]. The computer arts of Struycken [3], Brown [4] and evolutionary architecture of Frazer [5] are classical examples of CA based arts. Moreover, CA have been used for music composition, for example, Xenakis [6] and Miranda [7].

The main characteristics of CA that makes them particularly interesting to digital artists is their ability to produce interesting and very complex patterns on the basis of very simple rules. This fact has been noted by Wolfram, who himself produced some CA art 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 reminiscent of nature, but often unlike anything ever seen before" [8, p.11]. There have been some interesting attempts to develop means of controlling emergence of aesthetic behaviour in CA [9], [10], [11], [12] but with less success. Furthermore, it was observed that although classical one-dimensional CA with binary states can exhibit complex behaviour, multi-state 2D CA have significantly increased the complexity of behaviour, therefore, generating very complex symmetrical

patterns [13], [14] which are extremely challenging to generate using conventional mathematical methods.

Aesthetic judgements have long been hypothesised to be influenced by the stimulus degree of order (i.e. symmetry) and complexity. On this ground Birkhoff suggested a 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 [15].

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

Although the validity of Birkhoff's approach in penalising complexity has been challenged by empirical studies in [16], [17], the notions of order and complexity and objective methods to quantify them remains a prominent parameter in aesthetic evaluation functions.

Shannon's introduction of *information theory* provided a mathematical model to measure the degree of uncertainty (entropy) associated with a random variable [18]. Entropy then regarded as a measure of *order* and *complexity*. A low entropy implies low uncertainty so the message is highly predictable, ordered and less complex. And high entropy implies a high uncertainty, less predictability, highly disordered and complex. Moles [19], Bense [20] and Arnheim [21] 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. Since then entropy is commonly used to measure order and complexity in most of aesthetic evaluation functions [22], [23], [24], [25], however entropy fails to discriminate accurately structurally different patterns in two-dimensions [26].

In this paper, following our earlier works [14], [26], [27], we investigate the computational notion of aesthetic in CA as non-linear dynamical systems. Particularly, we examine *information gain* as a spatial complexity measure for discriminating 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 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 4 gives details of

experiments that test the effectiveness of information gain. The paper closes with a discussion and summary of findings.

II. CELLULAR AUTOMATA

Definition 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).
- $S = \{1, 2, \ldots, k\}$ 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 $\{e_a\}$, $a = 1, 2, \ldots, N$. The neighbourhood of cell $r = (i, j)$ is $\{r + e_1, r + e_2, \ldots, r + e_N\}$. A 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 $(s_1, s_2, \ldots, s_N)$. 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 $(s_1, s_2, \ldots, s_N)$ of cells in its neighbourhood: $s_1(t + 1) = f(s_1, s_2, \ldots, s_N)$. A quiescent state s_q satisfies $f(s_q, s_q, \ldots, s_q) = s_q$.

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 $c^0, c^1, c^2, \ldots, c^{T-1}$, 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 nonlinear 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.

III. INFORMATION GAIN MEASURE

The *information theory* was an attempt to address communication over an unreliable channel [18]. Entropy is the core of this theory [28]. 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:

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

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, $\log_2(|\mathcal{X}|)$:

$$0 \leqslant H \leqslant \log_2(|\mathcal{X}|) \qquad (3)$$

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. 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.

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. 1.

$$order \xleftrightarrow{\text{regular structure|irregular structure|structureless}} disorder$$

Fig. 1. The spectrum of spatial complexity.

A 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 [27].

A measure introduced in [29], [30], [31] 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,

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

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

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

where $P(x, y)$ is the joint probability, $\text{prob}(X = x, Y = y)$. $\overline{G}$ is also known as the conditional entropy, $H(X|Y)$ [28]. 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 $\overline{G}_{X,Y}$ are

$$0 \leqslant \overline{G}_{X,Y} \leqslant \log_2 |\mathcal{X}|. \qquad (6)$$

Definition 2: 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, $\overline{G}$ can be calculated by considering the distribution of cell states over pairs of cells r, s,

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

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

The vertical, horizontal, primary diagonal ($\searrow$) and secondary diagonal ($\nearrow$) 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 be 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
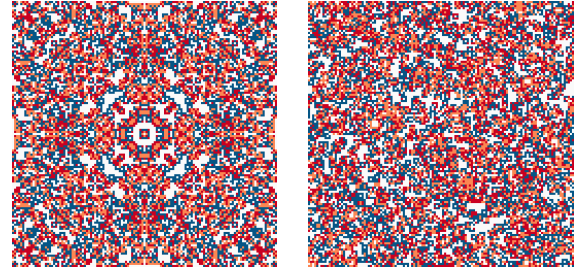
$$\Delta \overline{G}_{i,j\pm 1} = |\overline{G}_{i,j+1} - \overline{G}_{i,j-1}| \quad (9a)$$

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

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

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

The advantages of $\overline{G}$ over H in discriminating structurally different patterns is illustrated in Fig. 2 where the 4-state 2D CA configurations with the complete symmetrical (Fig. 2a), the partially structured (Fig. 2b) and the structureless and random (Fig. 2c) patterns are evaluated. 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 (i.e. symmetries) too. Fig. 2 clearly demonstrates the drawbacks of entropy to discriminate structurally different 2D patterns. In other words, entropy is invariant to spatial rearrangement of composing elements. This is in contrast to our intuitive perception of the complexity of patterns and is problematic for the purpose of measuring the order and complexity of multi-state 2D CA configurations for aesthetic evaluation.

(a)	(b)
$H = 1.98523$	$H = 1.98523$
$\overline{G}_{i,j+1} = 1.91810$	$\overline{G}_{i,j+1} = 1.92474$
$\overline{G}_{i,j-1} = 1.91810$	$\overline{G}_{i,j-1} = 1.92484$
$\Delta \overline{G}_{i,j\pm 1} = 0$	$\Delta \overline{G}_{i,j\pm 1} = 0.00010$
$\overline{G}_{i-1,j} = 1.91810$	$\overline{G}_{i-1,j} = 1.92617$
$\overline{G}_{i+1,j} = 1.91810$	$\overline{G}_{i+1,j} = 1.93635$
$\Delta \overline{G}_{i\pm 1,j} = 0$	$\Delta \overline{G}_{i\pm 1,j} = 0.00018$
$\overline{G}_{i-1,j+1} = 1.95562$	$\overline{G}_{i-1,j+1} = 1.95887$
$\overline{G}_{i+1,j-1} = 1.95562$	$\overline{G}_{i+1,j-1} = 1.95919$
$\Delta \overline{G}_{P_d} = 0$	$\Delta \overline{G}_{P_d} = 0.00032$
$\overline{G}_{i+1,j+1} = 1.95562$	$\overline{G}_{i+1,j+1} = 1.95570$
$\overline{G}_{i-1,j-1} = 1.95562$	$\overline{G}_{i-1,j-1} = 1.95560$
$\Delta \overline{G}_{S_d} = 0$	$\Delta \overline{G}_{S_d} = 0.00010$

(c)

$H = 1.98523$
$\overline{G}_{i,j+1} = 1.98474$
$\overline{G}_{i,j-1} = 1.98467$
$\Delta \overline{G}_{i,j\pm 1} = 0.00007$
$\overline{G}_{i-1,j} = 1.98499$
$\overline{G}_{i+1,j} = 1.98503$
$\Delta \overline{G}_{i\pm 1,j} = 0.00004$
$\overline{G}_{i-1,j+1} = 1.98514$
$\overline{G}_{i+1,j-1} = 1.98509$
$\Delta \overline{G}_{P_d} = 0.00005$
$\overline{G}_{i+1,j+1} = 1.98477$
$\overline{G}_{i-1,j-1} = 1.98465$
$\Delta \overline{G}_{S_d} = 0.00012$

Fig. 2. The comparison of H with measures of $\overline{G}_{i,j}$ for structurally different 4-state CA configurations.

IV. 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 I) maps four states, represented by *white*, *red*, *blue* and *orange*; the quiescent state is *white*. The experiments are conducted with four different ICs: (1) all quiescent states cells except for a single cell (Fig. 3a) (2) a right oriented 5 cell (Fig. 3b), (3) a left oriented 5 cell (Fig. 3c) and (4) a random configuration with 2112 *white*

TABLE I
UPDATE RULE OF EXPERIMENTAL CELLULAR AUTOMATON.

$L = 65 \times 65$ (4225 cells).
$S = \{0, 1, 2, 3\} \equiv \{\ \square,\ \blacksquare,\ \blacksquare,\ \blacksquare\ \}$
$\mathcal{N}$: Moore neighbourhood
$f : S^9 \mapsto S$

$$f(s_{i,j})(t) = s_{i,j}(t+1) = \begin{cases} 1 & \text{if } s_{(i,j)}(t) = 0 & \text{and } \sigma = 1 \\ 3 & \text{if } s_{(i,j)}(t) = 1 - 3 & \text{and } \sigma = 2 \\ 2 & \text{if } s_{(i,j)}(t) = 1 - 3 & \text{and } \sigma = 3 \\ 0 & \text{otherwise} \end{cases}$$

where σ is the sum total of the neighbourhood states.

quiescent states cells covering $\approx 50\%$ of the lattice, 749 *red*, 682 *blue* and 682 *orange* cells (Fig. 3d). The experimental rule has been iterated synchronously for 150 successive time steps. Figs. 4, 5, 6 and 7 illustrate a sample of time steps starting from four different ICs. Then the sequence of configurations are analysed by 9a, 9b, 9c and 9d.
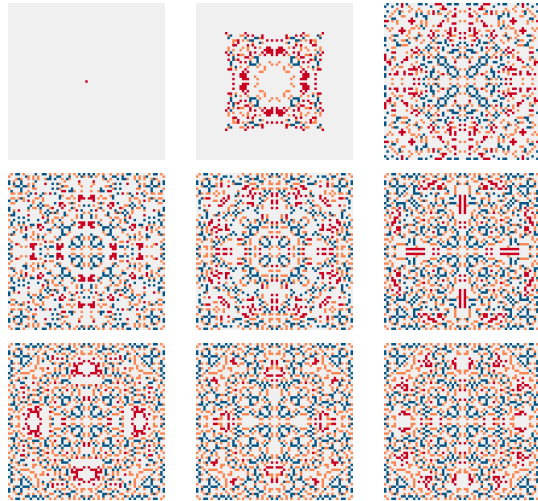

Fig. 3. The four different ICs.

Fig. 4. Space-time diagram of the experimental cellular automaton for sample time steps starting from the single cell IC (3a).

The behaviour of cellular automaton from the single cell IC is a sequence of symmetrical patterns (Fig. 4). This fact has been reflected on the measurements of $\Delta\overline{G}s$ (Fig. 8), where they are constant for the 150 time steps ($\Delta\overline{G}_{i,j\pm1} = \Delta\overline{G}_{i\pm1,j} = \Delta\overline{G}_{P_d} = \Delta\overline{G}_{S_d} = 0$). 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 (Fig. 14).
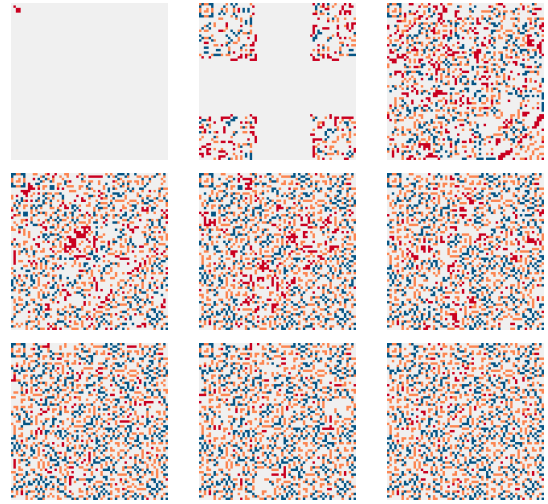

Fig. 5. Space-time diagram of the experimental cellular automaton for sample time steps starting from the 3b IC.
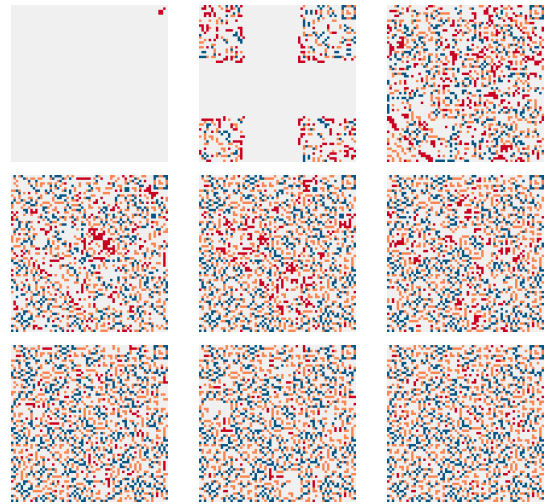

Fig. 6. Space-time diagram of the experimental cellular automaton for sample time steps starting from the 3c IC.

The behaviour of cellular automaton from two 5 cell ICs (3b and 3c) are sequence of symmetrical patterns with different orientations (Figs. 5, 6). The measurements of H for these two sequences of structurally different but symmetrical configurations are identical from $t = 0$ to $t = 150$, where $H_0^{3b} = H_0^{3c} = 0.01321$ and $H_{150}^{3b} = H_{150}^{3c} = 1.43241$ (Fig. 13). 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. 9, 10). This is further illustrated in Fig. 11 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 3b and 3c ICs.

The behaviour of cellular automaton from the random IC is a sequence of irregular structures (Fig. 7). The formation of patterns with local structures has reduced the values of $\Delta\overline{G}s$ until a stable oscillating pattern is attained (Fig. 14). This is an indicator of the development of irregular structures. However
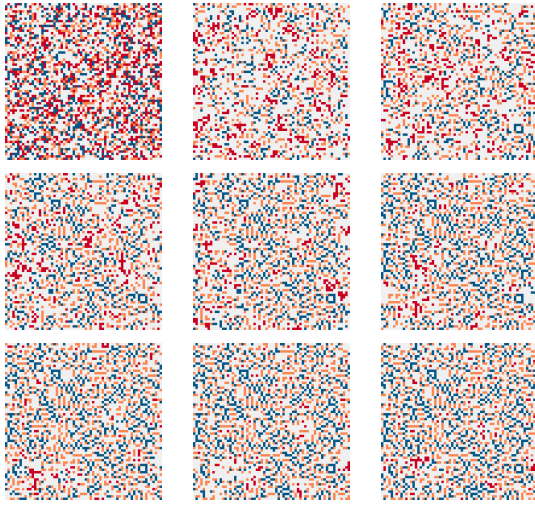
Fig. 7. Space-time diagram of the experimental cellular automaton for sample time steps starting from the random (3d) IC.

the patterns are not random patterns since the maximum four-state value $\log_2(4) = 2$ (Eq. 6).
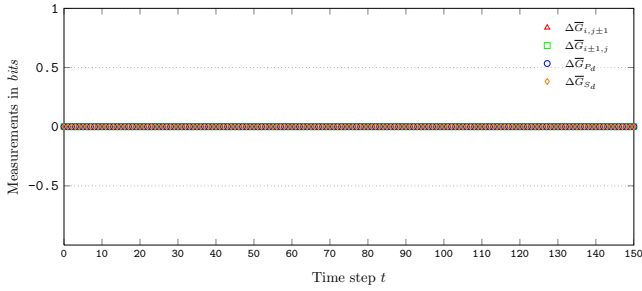

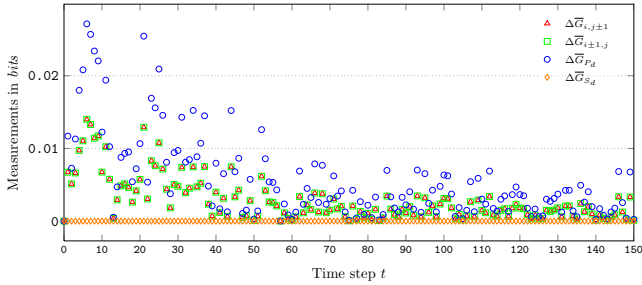Fig. 8. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 3a IC.

Fig. 9. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 3b IC.

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 to discriminate between the diversity of patterns that can be generated by various CA. The structured but asymmetrical patterns emerging from the random
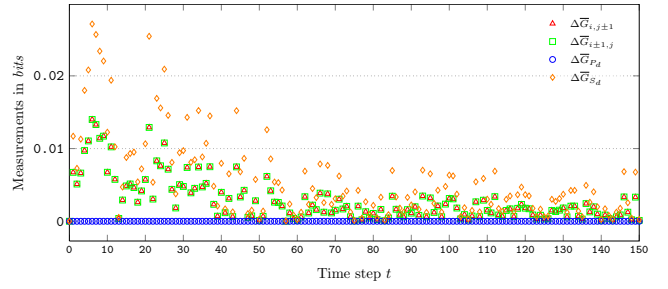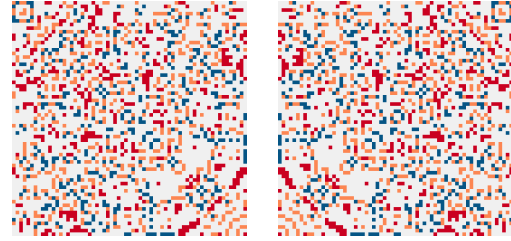

Fig. 10. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 3c IC.

(a)	(b)
$H = 1.42929$	$H = 1.42929$
$\overline{G}_{i,j+1} = 1.36140$	$\overline{G}_{i,j+1} = 1.36140$
$\overline{G}_{i,j-1} = 1.36538$	$\overline{G}_{i,j-1} = 1.36538$
$\Delta\overline{G}_{i,j\pm1} = 0.00398$	$\Delta\overline{G}_{i,j\pm1} = 0.00398$
$\overline{G}_{i-1,j} = 1.36140$	$\overline{G}_{i-1,j} = 1.36538$
$\overline{G}_{i+1,j} = 1.36538$	$\overline{G}_{i+1,j} = 1.36140$
$\Delta\overline{G}_{i\pm1,j} = 0.00398$	$\Delta\overline{G}_{i\pm1,j} = 0.00398$
$\overline{G}_{i-1,j+1} = 1.36634$	$\overline{G}_{i-1,j+1} = 1.37148$
$\overline{G}_{i+1,j-1} = 1.37431$	$\overline{G}_{i+1,j-1} = 1.37148$
$\Delta\overline{G}_{P_d} = 0.00797$	$\Delta\overline{G}_{P_d} = 0$
$\overline{G}_{i-1,j-1} = 1.37148$	$\overline{G}_{i-1,j-1} = 1.37431$
$\overline{G}_{i+1,j+1} = 1.37148$	$\overline{G}_{i+1,j+1} = 1.36634$
$\Delta\overline{G}_{S_d} = 0$	$\Delta\overline{G}_{S_d} = 0.00797$

Fig. 11. The comparison of H, $\overline{G}s$ and $\Delta\overline{G}s$ at $t = 40$ for 3b (a) and 3c (b) ICs.

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.
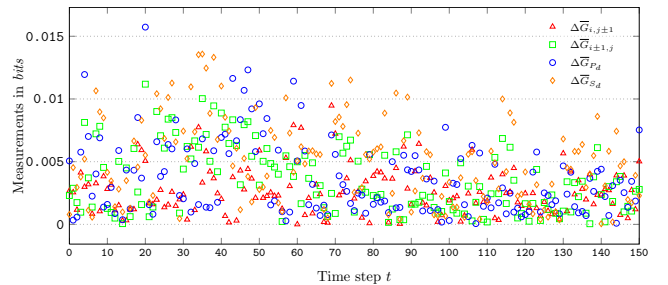

Fig. 12. The measurements of $\Delta\overline{G}s$ for 150 time steps starting from 3d IC.
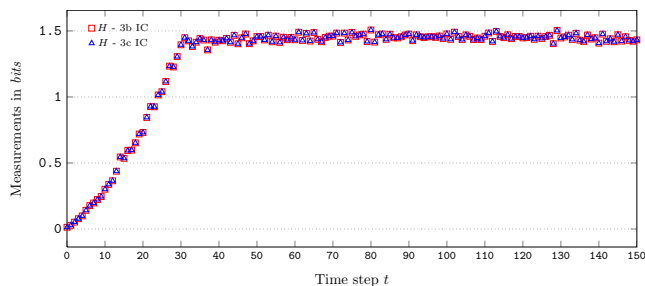
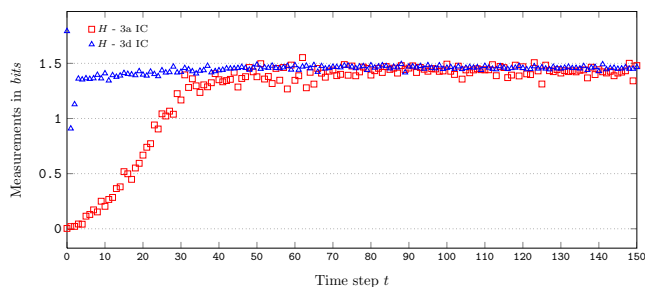Fig. 13.   The measurements of H starting from 3b and 3c ICs.

Fig. 14.   The measurements of H starting from 3a and 3d ICs.

V. CONCLUSIONS

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. This paper reports on a set of experiments for four different initial conditions of a cellular automaton. The potential of mean information gain for distinguishing multi-state 2D CA patterns is demonstrated. So the measure appears to be particular good at distinguishing different kinds of symmetries and random patterns from non-random patterns. 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

[1] K. C. Knowlton, "A computer technique for producing animated movies," in *Proceedings of the April 21-23, 1964, spring joint computer conference*.   ACM, 1964, pp. 67–87.

[2] L. Schwartz and L. Schwartz, *The Computer Artist's Handbook: Concepts, Techniques, and Applications*.   W W Norton & Company Incorporated, 1992.

[3] P. Struycken, "Splash 1972/1974," in *Artist and computer*, R. Leavitt, Ed.   Harmony Books, 1976, pp. 30–31.

[4] P. Brown, "Stepping stones in the mist," in *Creative evolutionary systems*.   Morgan Kaufmann Publishers Inc., 2001, pp. 387–407.

[5] J. Frazer, *An evolutionary architecture*.   Architectural Association Publications, Themes VII, 1995.

[6] I. Xenakis, *Formalized music: thought and mathematics in composition*. Pendragon Press, 1992.

[7] E. Miranda, *Composing Music with Computers*, ser. Composing Music with Computers.   Focal Press, 2001, no. v. 1.

[8] S. Wolfram, *A New Kind of Science*.   Wolfram Media Inc., 2002.

[9] W. Li, "Complex patterns generated by next nearest neighbors cellular automata," *Computers & Graphics*, vol. 13, no. 4, pp. 531–537, 1989.

[10] K. Sims, "Interactive evolution of dynamical systems," in *Toward a practice of autonomous systems: Proceedings of the first European conference on artificial life*, 1992, pp. 171–178.

[11] W. K. Mason, "Art from cellular automata and symmetrized dot-patterns," *Computers & graphics*, vol. 16, no. 4, pp. 439–441, 1993.

[12] D. Ashlock and J. Tsang, "Evolved art via control of cellular automata," in *Evolutionary Computation, 2009. CEC'09. IEEE Congress on*.   IEEE, 2009, pp. 3338–3344.

[13] M. A. Javaheri Javid and R. te Boekhorst, "Cell Dormancy in Cellular Automata," in *International Conference on Computational Science (3)*, ser. Lecture Notes in Computer Science, V. N. Alexandrov, G. D. van Albada, P. M. A. Sloot, and J. Dongarra, Eds., vol. 3993.   Springer, 2006, pp. 367–374.

[14] M. A. Javaheri Javid, M. M. al Rifaie, and R. Zimmer, "Detecting Symmetry in Cellular Automata Generated Patterns Using Swarm Intelligence," in *Theory and Practice of Natural Computing*, ser. Lecture Notes in Computer Science, A.-H. Dediu, M. Lozano, and C. Martín-Vide, Eds., vol. 8890.   Springer International Publishing, 2014, pp. 83–94.

[15] G. Birkhoff, *Aesthetic Measure*.   Harvard University Press, 1933.

[16] D. J. Wilson, "An experimental investigation of Birkhoff's aesthetic measure," *The Journal of Abnormal and Social Psychology*, vol. 34, no. 3, p. 390, July 1939.

[17] H. J. Eysenck, "The empirical determination of an aesthetic formula." *Psychological Review*, vol. 48, no. 1, p. 83, 1941.

[18] C. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423 & 623–656, Oct. 1948.

[19] A. Moles, *Information theory and esthetic perception. Trans. JE Cohen*. U. Illinois Press, 1968.

[20] M. Bense, *Kleine abstrakte ästhetik [small abstract aesthetics]*, ser. Edition Rot.   E. Walther, March 1969, vol. 38.

[21] R. Arnheim, "Towards a psychology of art/entropy and art an essay on disorder and order," *The Regents of the University of California*, 1966.

[22] H. W. Franke, "A cybernetic approach to aesthetics," *Leonardo*, vol. 10, no. 3, pp. 203–206, 1977.

[23] P. Machado and A. Cardoso, "Computing aesthetics," in *Advances in Artificial Intelligence*.   Springer, 1998, pp. 219–228.

[24] T. Staudek, "Exact Aesthetics. Object and Scene to Message," Ph.D. dissertation, Faculty of Informatics, Masaryk University of Brno, 2002.

[25] J. Rigau, M. Feixas, and M. Sbert, "Informational aesthetics measures," *Computer Graphics and Applications, IEEE*, vol. 28, no. 2, pp. 24–34, 2008.

[26] M. A. Javaheri Javid, M. M. al Rifaie, and R. Zimmer, "An Informational Model for Cellular Automata Aesthetic Measure," in *AISB 2015 Symposium on Computational Creativity*, University of Kent, Canterbury, UK, 2015.

[27] M. A. Javaheri Javid, T. Blackwell, R. Zimmer, and M. M. al Rifaie, "Spatial Complexity Measure for Characterising Cellular Automata Generated 2D Patterns," in *Progress in Artificial Intelligence, 17th Portuguese Conference on Artificial Intelligence, EPIA 2015, Coimbra, Portugal, September 8-11, 2015, Proceedings.*, ser. Lecture Notes in Artificial Intelligence, Francisco Pereira, Penousal Machado, Ernesto Costa and Amílcar Cardoso, Ed., vol. 9273.   Springer International Publishing, 2015.

[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*.   Wiley-Interscience, 2006.

[29] J. E. Bates and H. K. Shepard, "Measuring complexity using information fluctuation," *Physics Letters A*, vol. 172, no. 6, pp. 416–425, 1993.

[30] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, and H. Schein-graber, "A comparative classification of complexity measures," *Chaos, Solitons & Fractals*, vol. 4, no. 1, pp. 133–173, 1994.

[31] Andrienko, Yu. A., Brilliantov, N. V., and Kurths, J., "Complexity of two-dimensional patterns," *Eur. Phys. J. B*, vol. 15, no. 3, pp. 539–546, 2000.