An Approach to Searching for Two-Dimensional Cellular Automata for Recognition of Handwritten Digits

C.C. Oliveira Jr.[2] and P.P.B. de Oliveira[1,2]

Universidade Presbiteriana Mackenzie
[1] Faculdade de Computação e Informática
[2] Pós-Graduação em Engenharia Elétrica
Rua da Consolação 896 - Consolação
01302-907 São Paulo, SP – Brazil
carloscandidojr@gmail.com,
pedrob@mackenzie.br

Abstract. One of the contexts in which cellular automata have clearly demonstrated their effectiveness has been in problems involving strong and explicit spatial constraints, as happens in pattern formation and growth. By analogy, attempts to use cellular automata in pattern recognition have also been used in the literature and some progress has been made. However, in general, they still represent more of an unfulfilled promise, due to the lack of a recognition model which cellular automata would naturally fit in, the lack of effective ways to implement it, and the lack of generality of the available approaches. Here, experimental results are reported in the direction of using cellular automata in the task of handwritten digit recognition, in which an evolutionary algorithm searches for two-dimensional cellular automata rules that would transform a given digit image into a match, as close as possible, to a prototype image of that family, so that, the closer the match, the better the recognition of the input image. Although the results reported might still fall shorter than consolidated commercial techniques for the task, the approach presented is quite attractive in terms of the efficacy level it allowed to achieve, and because of its simplicity, which suggests a potential generality from the perspective of its use in other domains.

Keywords: Cellular automaton, pattern recognition, handwritten digit, handwritten character, evolutionary computation, genetic algorithm.

1 Introduction

Cellular automata (CAs) are distributed systems comprising a number of typically identical simple components, with local connectivity over a regular lattice, whose global configuration changes over time, according to a local state transition rule [17]. Quite remarkable about them is that they can be regarded both as an universal computing device, and as a dynamical system, discrete in space, time and state variables [17].

A. Gelbukh and E.F. Morales (Eds.): MICAI 2008, LNAI 5317, pp. 462–471, 2008.

One of the contexts in which cellular automata have clearly demonstrated their effectiveness has been in problems involving strong and explicit spatial constraints, as happens in the growth of urban settlements [14]; in the context of natural hazards in the environment due to the dispersion of, say, forest fire [25] or volcanic lava [27]; propagation of contagious diseases [26]; and pattern formation in biology [24] or engineering [21, 22]. And based upon their computing abilities, one might wonder, analogously, about using them also as pattern recognisers of spatial data.

Attempts to use cellular automata in pattern recognition in general, mainly spatial constrained data, have been a recurring theme in the literature, so that some progress has indeed been made [2, 3, 5, 8, 19, 20]. However, in general, they still fall short of the requirements for real-world, effective applications, thereby constituting more of an unfulfilled promise. These shortcomings have been due to the lack of a recognition model which cellular automata would naturally fit in, the lack of effective ways to implement it, and the lack of generality of the available approaches. Symptomatic of this situation is, for instance, the fact that, out of the 117 citations in [28], and out of the 255 citations in [29], none of them refer to cellular automata; also, out of the more than 88,000 references in [16], the joint search for ´cellular automata´ and ´cellular automaton´ entries return only 281 hits from the database, a mere 0.3%.

Recognition of handwritten digits by computational systems is a non-trivial task, due to the huge diversity of forms characters may display. In fact, the strong spatial constraints embedded in them are essential and inherent to their nature, so that they should thus be accounted for by any recognition system. Despite this difficulty, the task has been the subject of extensive studies, to the extent that it can be considered a well solved task presently, with even very effective commercial products available, based on consolidated techniques; nevertheless, contrarily to what one might think, new techniques continue to be proposed, with hopes for increased performance and generality [6, 12, 13, 18].

All this context renders the recognition of handwritten digits quite appealing and challenging for a cellular automaton based approach, which is precisely the motivation of the present work; here, an evolutionary algorithm searches for two-dimensional cellular automata rules that would transform a given input digit image into a match, as close as possible, to a prototype image of that family, so that, the closer the match, the better the recognition of the input image.

Although the achievements reported might still fall shorter than other existing competing models for the task, the computational technique presented is quite attractive not only in terms of the efficacy level it allowed to achieve, but also because of its simplicity, which suggests a potential generality from the perspective of its use in other domains.

In the next section the recognition approach at issue is introduced. Section 3 then describes the experiments performed and Section 4 the results obtained. The paper closes with Section 5, with an overall appreciation of the work, including a discussion on its roots and its natural follow-up.

2 The Approach

2.1 Cellular Automata as Dynamical Processors of Spatial Patterns

Cellular automata (CAs) are distributed systems comprising a large number of identical simple components with local connectivity [17]. These structures, the cells, form an n-dimensional grid, and can take on any of k possible states that transform over time according to a local rule or function. This rule determines the state of every cell at instant t+1 according to the states of its neighbouring cells at time t. The number of cells in the neighbourhood is usually determined by the radius r of the rule, which is the distance between the cell at issue and the farthest cell in its neighbourhood.

For given values of k and r, the number of possible rules defines the rule space, where such rules are usually enumerated according to Wolfram´s lexicographical order [17]; in the case of binary CAs, the latter is such that the output bits of each rule are arranged from neighbourhood 111…1 to 000…0. For instance, the rule space of the ´elementary´ CAs is formed by all 256 possible rules with k=2 and r=1, its rule 232 being represented by the binary number 11101000 formed out of its output bits, as illustrated in Figure 1, where whites and blacks represent 0s and 1s, respectively.

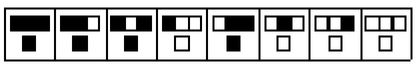

Fig. 1. Representation of rule 232 of the elementary space

A two-dimensional cellular automaton consists of a two-dimensional grid, with state transitions defined according to a two-dimensional neighbourhood. The most widely used neighbourhoods are Moore neighbourhood, that considers a cell and all its eight immediate neighbours, and von Neumann neighbourhood, which incorporates only a cell and its four immediate vertical and horizontal neighbours.

In successful approaches to the problem of handwritten digit recognition, computational systems typically allow input of a standard form and subsequently compare a new input with a standard prototype, explicitly or implicitly acquired, through learning, in a previous stage [1].

Likewise, here the handwritten digits are represented as a two-dimensional, binary cellular automaton initial configuration, with black pixels standing for 1 and white pixels for 0. A CA rule is then applied that leads the initial configurations representing digits, to corresponding final configurations, associated to the classes of digits, from 0 to 9.

All experiments reported below relied upon input digits comprising a set of well-known images of handwritten digits in the literature [10], as exemplified in Figure 2. Handwritten digit families were then defined, containing groups of handwritten digits representing the same digit, and the families organised in distinct sizes, according to the number of images within it. For example, a digit family for digit 3 and size 5 is a group formed by 5 images of digit 3, as depicted in Figure 3.

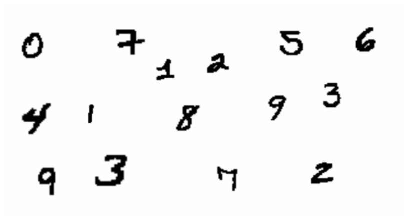

Fig. 2. A sample of initial configurations

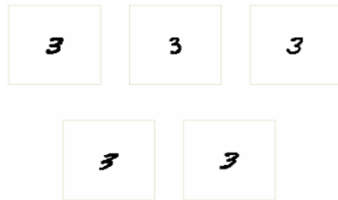

Fig. 3. A digit family of size 5 for digit 3

2.2 The Recognition Model

A key notion in the recognition model developed is that of a *prototype*, which is defined as a particular image that might represent all the images in a digit family. In order to calculate the prototype image, it suffices to calculate the mean image within a group, so that each point in the prototype represents the arithmetic mean of the corresponding points across all images at issue.

As suggested earlier, recognition is deemed effective when, upon applying a given CA rule to initial configurations representing the same digit, the final configurations achieved by the CA lattice preserve some similarity with each other, according to the digit family the input belongs to. Such similarity is verified by calculating the distance between the final configuration generated from a given input digit, and the prototype of the final configurations that has been previously calculated for the handwritten digit family of the input digit. Naturally, ten prototypes needed to be calculated, one for each digit.

Empirical trials in the problem led to Euclidean distance being employed as the basis for the similarity distances between the images, mainly because of its widespread use.

Recognition of a specific handwritten digit is then granted if its distance to the prototype of the family (or class) it is supposed to belong to is the smallest. This means that the cellular automaton rule employed to process the given initial configuration (the input handwritten digit) has led to a final configuration that is close, in Euclidean distance, to the prototype of the family the input digit belongs to.

For the generation of the prototypes, an image set was established, the ´training´ set; analogously, a test set of images was also defined, consisting of those used to verify the effectiveness of the approach. The creation of the training and test sets relied upon 1500 images that were initially selected from [10], with the constraint of their not being terribly ambiguous to humans. Then, 10% of them were drawn to form the test set, 15 for each digit class. The remaining subset of 1350 images was further subdivided into 5 groups, the first one having 270 images and the subsequent ones obtained by adding 270 new images to the preceding one.

Various recognition trials were carried out with the previously defined image sets; however, considering that each new sub-group of the training set that was added to the previous only incremented recognition slightly, the smaller group of 270 images (27 from each digit class) ended up being employed for training in the experiments reported below, with the gain of a much smaller computation time.

3 Experiments and Results

3.1 Recognition with Rules Enumerated through Crossed ECAs

The two-dimensional CA space delimited by Moore neighbourhood was defined as the target of the searches to be carried out. This choice derived from the huge range of possibilities it supports, comprising a total of 2^{512} rules, that even includes the other natural candidate, von Neumann neighbourhood space.

However, before embarking on the evolutionary searches in the chosen target CA rule space, an initial enumerative experiment was carried out in a much smaller space, that can be shown to constitute a subspace of Moore neighbourhood, the one that is brought about by what is named here the ´crossed elementary cellular automata´ (Crossed ECAs).

Crossed ECAs entail application of two-dimensional cellular automata rules, through the use of the application of elementary CA rules, in two stages, as follows:

1. An elementary rule is applied to the rows of the two-dimensional lattice;
2. Another elementary rule (possibly the same one as before) is applied to the columns of the lattice, using the outcome of the previous stage as input.

In the application of ECA rules to rows and columns at the two stages, the two-dimensional cellular automaton is partitioned and considered a one-dimensional elementary cellular automaton with periodic boundary conditions. Each application (rows and columns) is then considered a separate CA iteration.

The advantage of such a scheme is that it samples a meaningful subset of all possible two-dimensional CAs of interest, in that it allows transfer of knowledge of elementary CAs to Moore neighbourhood CAs. To enable this, a set of all ordered ECA rule pairs was formed consisting of the pairwise grouping of elementary rules, excluding those classified as complex (i.e., edge-of-chaos) or chaotic, according to Li-Packard classification [11]. These exclusions, associated to the individual rules´ dynamical behaviours aimed at preventing the effect of an excessive convergence time in the temporal evolutions, or even the lack of convergence at all.

The ordered pairs were obtained by enumeratively combining the individual elementary CA rules, one by one, without repeating rules, resulting in 5329 rule pairs; each one was then applied to a handwritten digit following the Crossed ECA structure.

The 5329 rule pairs were then submitted to the recognition process, according to the following sequence: generation of the 10 prototypes, by applying the rule pair to the training set; generation of the final configuration of each image in the test set, by applying the rule pair for a certain number of time steps; calculation of the Euclidian distances between each final configuration and the 10 prototypes; and, verification of the minimum distances, so that a given input configuration was granted recognised if its minimum distance corresponded exactly to the expected prototype for it.

After running the process for all 5329 rule pairs, the optimal rule pair 62-168 recognised 95.3% of the entire test set, much higher than the average of the set. This corresponded to recognition of 143 images from the test set, the pair having reached such a recognition rate by evolving the initial configuration for 2 iterations. Figure 4 shows the prototypes generated by that rule pair. It is interesting to realise that the

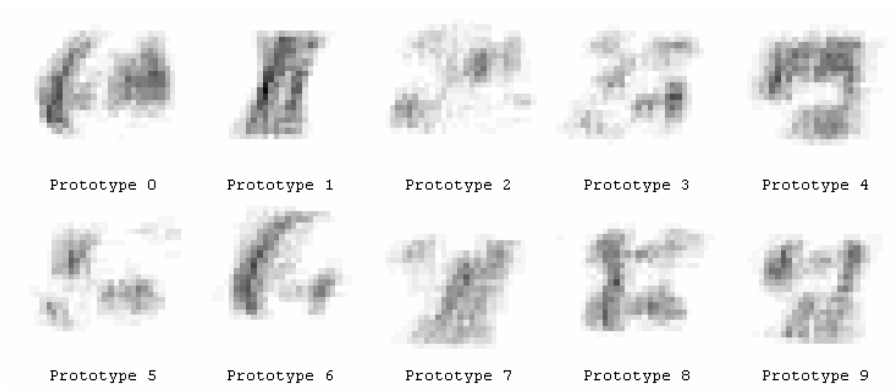

Fig. 4. Prototypes generated by the rule pair 62-168, each one corresponding to a digit family

generated prototypes have very little relation to the overall shape of each digit, even considering only 2 time steps; however, looking at them in more detail it seems that they indeed represent, even though vaguely, some key portion of each digit, as can be noticed, for instance, comparing the prototypes for digits 6 and 9, or those for digits 1 and 7. This strongly suggests that the rules seem to be doing a preprocessing of the input digit, so that they would become more amenable for recognition. But what is remarkable is that such a feature was not embedded in the system at all, explicitly or otherwise.

3.2 Recognition with Rules Searched by a Genetic Algorithm

Since the entire Crossed ECA space is rather small, composed of only $256^2=2^{16}$ rules, and the biased, enumerative experiment above showed that even in this space good rules do exist, better rules are likely to exist within the much larger Moore neighbourhood rule space.

In order to search for rules in this vast space, a genetic algorithm (GA) was developed, based on the one described in [9]. In the algorithm, each individual is the binary, 512 bit-long representation of one possible rule in Moore neighbourhood space. The representation followed Wolfram´s lexicographical ordering mentioned above, now applied in two dimensions.

The fitness evaluation of each candidate rule relies upon the same process described in the previous experiment, which considers the images in the training set. The actual value of the fitness of a candidate rule is then the recognition rate of the rule, over the images in the test set.

GA runs were devised so as to try a number of different parameters. Among all the runs, the optimal rule recognised 98% of the images in the test set, thus representing an improvement on the best rule pair of the enumerative experiment. This result was obtained using the following parameters: population size of 50 individuals (rules); single-point crossover, at 100% rate; mutation rate of 0.02% per bit; selection of 20% of the elite, the remaining 80% of the rules being randomly chosen from the entire population; CAs applied for 2 iterations; and 120 generations per run.

Recognition of 98% of the test set implied that only three digits were unrecognised, the ones depicted in Figure 5. Inspection of the unrecognised digits did not suggest any overall reason related to their recognition failure; however, it should be noticed that the unrecognised digit 2 is dubious even for human recognisers. Nevertheless, in view of the vast rule space concerned, a rule may exist that might allow a further increase in the recognition rate of the test set, maybe even leading to full recognition.

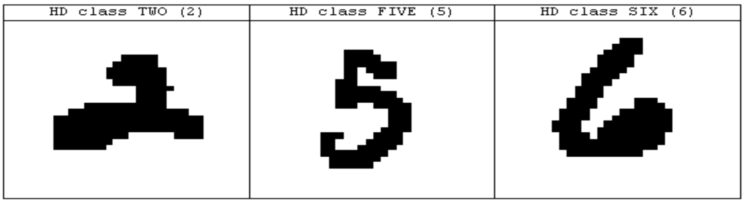

Fig. 5. Handwritten digits not recognised by the best rule found by the GA, and their respective digit family

For the purposes of evaluating the effectiveness of the results obtained, two baseline or reference experiments were carried out. The first one derived from the idea of investigating CA-based recognition from a spectral context, that is, to check the effect of handling images in the frequency domain, chiefly due to their non-susceptibility to translation and other transformations. To this end, all images underwent discrete Fourier transform, and, in order to quantitatively ascertain similarity between the images, Euclidian distance was used between the magnitude of their components, referred to as ´spectral similarity´. Although a range of evaluation approaches for the rules were tried out, the best result reached only 82% recognition efficacy.

In the second baseline experiment, CAs were not employed at all to transform the initial configurations. Accordingly, the prototypes were generated directly from the images of the training set, and the test set was then checked by calculating the Euclidian distance of each image and its respective prototype. Hence, the key difference of this experiment was that the handwritten digit images did not undergo the transformations entailed by the CA rule application. As a result, 88.7% of the test set was recognised, a surprisingly high value that suggests the effectiveness of the notion of prototype, as used here.

Altogether, the baselines demonstrate the improved performance entailed by the CA rule image transformation, linked to the notion of (non-spectral) prototype. Although such comparisons might lag behind of other recognition techniques, particularly those in commercial systems, they certainly hint at the strengths of the proposed approach.

4 Concluding Remarks

Various CA based efforts in image processing, particularly those in edge detection, are available in the literature [4, 15, 30, 31, 32] and, naturally, CA based recognition approaches based on edge detection as well [2, 33, 34]. However, while this type of processing is explicitly targeted in the latter endeavours, here the CA rule being searched for has no ´obligation´ whatsoever of carrying out edge detection at any stage of its operation on the initial configuration. In fact, quite peculiar of the approach explored here is the total absence of a priori assumptions regarding the prototype nature, which does not have to abide by any predefined constraint. This departs this work the most from other related efforts. This might sound naïve at first glance, but this simple idea reflects the very essence of cellular automata nature, and, as such, the most natural way, we believe, towards their usage.

Such simplicity relies upon the notion of prototype (drawn from a rule's capacity to yield similar temporal evolutions to similar initial configurations of the same digit family), the reliance on a direct (non-spectral) approach for its creation, and use of plain Euclidean distance to drive the evaluation of a candidate rule. Although the present results may not have yet been entirely satisfactory, the conceptual structure stemmed from the work stands out in itself, both in effectiveness and simplicity, and paving the way for further enhancements.

The enumerative experiment relying on the elementary rule combinations showed to be relevant, in that the dynamical nature of the individual rules involved pointed at fruitful points in Moore neighbourhood space, as reflected by the good recognition efficacy achieved by rule pair 62-168. This realisation encouraged the search for rules in the complete space delimited by the Moore neighbourhood.

One possibility for further improvements in the efficacy level in the digit recognition task might be the idea of favouring the search towards rules that would be able to recognise simpler structures first, such as digits formed by straight lines only (namely, 1, 4 and 7) versus those fully rounded (i.e., digits 0, 3, 6, 8 and 9), or simple straight line segments versus curved segments. This might help the search to first tackle ´simpler´ problems, for only then to go freer; this is analogous to what has recently been done in [23], which was key to the results reported therein for the density and the parity problems, the best ones presently available for these tasks.

Further possible efficacy improvements might come from using much more sophisticated evolutionary algorithms than the very simple one described here; once again, [23] provides several good possible extensions. Quite encouraging in this respect is that, even with a simple evolutionary technique, good rules could be found.

A key issue in further developments of this work is to probe the scalability of the method, something that has not yet been looked at. Similarly, care should also be taken for the analysis of the unrecognised digits, since, as shown above, one of the test instances we used can be misleading even for human recognisers. Also, a proper, systematic comparison of the approach with existing consolidated techniques, even commercial ones, is tempting in the horizon, provided the required resources are made available.

Further topics suggestive of a distinct stance of investigation would include the use of non-uniform CAs, as in [19], in which distinct cells would be allowed to embed different rules, thus giving more flexibility for evolution to work, and the development of new applications, such as in recognition of other types of characters (letters or ideograms), thus allowing to probe the generality of the method, a feature that is certainly facilitated by its conceptual simplicity.

These are all appealing targets to be pursued in parallel and in sequence, whose ultimate goal is to put under scrutiny the usefulness of cellular automata as pattern recognisers, possibly allowing them to re-enact, in the new role, their quintessential, well established vocation as pattern generators.

Acknowledgement

We are very grateful to MackPesquisa – Fundo Mackenzie de Pesquisa, for various grants, in particular, the most recent one, from Edital 2007; Wolfram Research, for a Mathematica Academic Grant (No 1149); and FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo, for the research grant Proc. 2005/04696-3.

References

1. Bahlmann, C., Burkhardt, H.: The writer independent online handwriting recognition system frog on hand and cluster generative statistical dynamic time warping. IEEE Transactions on Pattern Analysis and Machine Intelligence 6(3), 299–310 (2004)
2. Cagnoni, S., Mordonini, M.: Cellular automata and computer vision: Some case studies. AI*IA Notizie XIV(2), 23–31 (2001)
3. Raghavan, R.: Cellular automata in pattern recognition. Information Sciences 70(1-2), 145–177 (1993)
4. Popovici, A., Popovici, D.: Cellular automata in image processing. In: Proc. of the 15th Int. Symp. on the Mathematical Theory of Networks and Systems, Univ. of Notre Dame (2002)
5. Ganguly, N., Das, A., Maji, P., Sikdar, B.K., Pal Chaudhuri, P.: Theory and application of cellular automata for pattern classification. Fundamenta Informaticae 58(3-4), 321–354 (2003)
6. Kherallah, M., Haddad, L., Alimi, A.M., Mitiche, A.: On-line handwritten digit recognition based on trajectory and velocity modeling. Pattern Recognition Letters 29(5), 580–594 (2008)
7. Guan, S., Zhang, S.: An evolutionary approach to the design of controllable cellular automata structure for random number generation. IEEE Transactions on Evolutionary Computation 7(1), 23–26 (2003)
8. Hogg, T., Huberman, B.: Parallel computing structures capable of flexible associations and recognition of fuzzy inputs. Journal of Statistical Physics, 41–115 (1985)
9. Jacob, C.: Illustrating Evolutionary Computation with Mathematica. Morgan Kaufmann, San Francisco (2001)
10. Lecun, Y., Corte, C.: The MNIST database of handwritten digits (last access, April 2008), `http://yann.lecun.com/exdb/mnist`
11. Li, W., Packard, N.: The structure of elementary cellular automata rule space. Complex Systems (4), 281–297 (1990)
12. Zhang, P., Bui, T.D., Suen, C.Y.: A novel cascade ensemble classifier system with a high recognition performance on handwritten digits. Pattern Recognition 40(12), 3415–3429 (2007)
13. Lauer, F., Suen, C.Y., Bloch, G.: A trainable feature extractor for handwritten digit recognition. Pattern Recognition 40(6), 1816–1824 (2007)
14. Pinto, N.N., Antunes, A.P.: Cellular automata and urban studies: A literature survey. ACE: Architecture, City and Environment 1(3), 368–399 (2007)
15. Rosin, P.L.: Training cellular automata for image processing. In: Kalviainen, H., Parkkinen, J., Kaarna, A. (eds.) SCIA 2005. LNCS, vol. 3540, pp. 195–204. Springer, Heidelberg (2005)

16. Price, K.: The annotated computer vision bibliography (last access, April 2008), http://www.visionbib.com
17. Wolfram, S.: A New Kind of Science, 1st edn. Wolfram Media (2002)
18. Savas, B., Eldén, L.: Handwritten digit classification using higher order singular value decomposition. Pattern Recognition 40(3), 993–1003 (2007)
19. Maji, P., Chaudhuri, P.P.: Non-uniform cellular automata based associative memory: Evolutionary design and basins of attraction. Information sciences 178(10), 2315–2336 (2008)
20. Maji, P., Chaudhuri, P.P.: Fault diagnosis of electronic circuits using cellular automata based pattern classifier. In: Prasad, B. (ed.) Soft Computing Applications in Industry, pp. 225–246. Springer, Berlin (2008)
21. Pikula, T., Gramatová, E., Fischerová, M.: Automatic design of cellular automata for generating deterministic test patterns. In: Digest of Papers: IEEE European Test Workshop (ETW), Maastricht, The Netherlands, pp. 285–286 (2003)
22. Dasgupta, P., Chattopadhyay, S., Chaudhuri, P.P., Sengupta, I.: Cellular automata-based recursive pseudoexhaustive test pattern generator. IEEE Transactions on Computers 50(2), 177–185 (2001)
23. Wolz, D., Oliveira, P.P.B.: Very effective evolutionary techniques for searching cellular automata rule spaces. Journal of Cellular Automata (forthcoming, 2008)
24. Deutsch, A., Dormann, S.: Cellular automaton modeling of biological pattern formation: Characterization, applications, and analysis. Birkhäuser, Boston (2005)
25. Encinas, A.H., Encinas, L.H., White, S.H., Rey, A.M., Sánchez, G.R.: Simulation of forest fire fronts using cellular automata. Advances in Engineering Software 38(6), 372–378 (2007)
26. White, S.H., Rey, A.M., Sánchez, G.R.: Modeling epidemics using cellular automata. Applied Mathematics and Computation 186(1), 193–202 (2007)
27. D'Ambrosio, D., Spataro, W.: Parallel evolutionary modelling of geological processes. Parallel Computing 33(3), 186–212 (2007)
28. Koerich, A.L., Sabourin, R., Suen, C.Y.: Large vocabulary off-line handwriting recognition: A survey. Pattern Analysis & Applications 6(2), 97–121 (2003)
29. Plamondon, R., Srihari, S.N.: Online and off-line handwriting recognition: a comprehensive survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(1), 63–84 (2000)
30. Batouche, M., Meshoul, S., Abbassene, A.: On solving edge detection by emergence. In: Ali, M., Dapoigny, R. (eds.) IEA/AIE 2006. LNCS (LNAI), vol. 4031, pp. 800–808. Springer, Heidelberg (2006)
31. Goi, H.T., Perkowski, M.: An original method of edge detection based on cellular automata (last access, 2008), http://web.cecs.pdx.edu/~mperkows/temp/June16/A-New-Method-of-Edge-Detection-Based-on-Cellular-Automaton-Algorithm.doc
32. Slatnia, S., Batouche, M., Melkemi, K.E.: Evolutionary cellular automata based-approach for edge detection. In: Masulli, F., Mitra, S., Pasi, G. (eds.) WILF 2007. LNCS (LNAI), vol. 4578, pp. 404–411. Springer, Heidelberg (2007)
33. Namvong, A., Wongthanavasu, S.: Facial features localization from range images using cellular automata model. In: Proc. of IAIT 2007: Int. Conf. on Advances on Information Technology 2007, Bangkok, Thailand (2007)
34. Wongthanavasu, S., Tangvoraphonkchai, V.: Cellular automata-based algorithm and its application in medical image processing. In: Proc. of ICIP 2007: IEEE International Conference on Image Processing, vol. 3, pp. III.41–III.44 (2007)