A Set of Spatial Characteristics in Cellular Automata Land Cover Change Modeling

Alexander ZAMYATIN

Abstract

An approach to the enhanced land cover change modeling using cellular automata and a set of 7 spatial characteristics that are often used for different geoinformatics problems is proposed. The informativeness of every characteristic in different cellular automata neighborhood is defined. An evaluation procedure of the characteristics' level of significance for its various combinations in a set is proposed for cellular automata transition rules definition. According to the evaluation procedure the significance levels of characteristics are defined.

1 Introduction

One of the key factors that influence the adequacy of land cover change cellular automata (CA) modeling is the determination of transition rules in each specific case (LI & YEH 2004). Nowadays there are a lot of different approaches for spatial characteristics used in various geoinformatics problems. For instance, in VERBURG et al. 2003 the enrichment factor for land use pattern analysis was used. In TISCHENDORF 2001 a set of landscape indices for landscape structure investigation was applied. Moreover, a preliminary research (ZAMYATIN 2007) shows promising results to use spatial characteristics for enhanced CA transition rules determination. A practical use of spatial characteristics to consider landscape spatial features more effectively, particularly for CA modeling, has some limitations – different informativeness of characteristics in different neighborhood size, ambiguousness of an appropriate set size as well as possible variants of its combination in the set in each specific case. As a result only few examples of using spatial features have been presented. For instance, in the work of BENENSON & TORRENS 2004 – the only spatial characteristic, in ZAMYATIN et al. 2007 – four spatial characteristics were used for land cover change modeling.

This work is aimed to enhance CA land cover change modeling using the extended set with 7 spatial characteristics that are often used in different geoinformatics areas, and to analyze these characteristics in a varying neighborhood size and determine their effective compatibility in a joint use.

2 CA Transition Rules in Modeling Process

ZAMYATIN et al. 2007 offer to determine CA transition rules in the land cover change modeling algorithm in every point of the analyzed area on the basis of transition of i-th to j-th land cover type as $p_{ij}^{res} \sim |1-(1+ p_{ij}^{prob})\cdot(1+ p_{ij}^{sp})\cdot(1+ p_{ij}^{add})|$, where p_{ij}^{prob} –probabilistic characteristic on the basis of a stochastic matrix, p_{ij}^{add} – probability on the basis of priory information from suitability maps (buffer zone, distance to roads, etc.). Probability p_{ij}^{sp} in the p_{ij}^{res} expression is proposed to define with the extended set of 7 different spatial characteristics: C_{FO} – frequency of occurrence, C_{LF} – level of fragmentation, C_{AFM} – average fractal measure of perimeters, C_{ADP} – average distance between patches, C_{LD} – level of division, C_{LPS} – level of patch size, C_{AFSPA} – average fractal size of perimeters-areas. Spatial characteristics are accounted for each i-th element of the image in a sliding window with $(2d+1) \times (2d+1)$ elements in a neighborhood, where d can be an order of the neighborhood using such features as the number of patches of all types, the number of type elements in the neighborhood, the patch number, etc.

In fact, spatial characteristics for land cover change modeling with CA could be used in many different ways. In this research it is proposed to calculate the p^{sp} probability for every i-th point of the image on the basis of the expression $p^{\mathrm{sp}} = |1 - d(\mathbf{F}_{\mathbf{Cloc}}, \mathbf{F}_{\mathbf{Cglb}})|$, where $d(\mathbf{F}_{\mathbf{Cloc}}, \mathbf{F}_{\mathbf{Cglb}})$ – Euclidean distance between vectors of the so-called «local» and «global» characteristics. Comparing these vectors for every landscape type in pairs we can define how similar both the spatial characteristics of a k-th type in the current sliding window and the average spatial characteristics of the whole image are. It allows to define the transition probability for every i-th central element of the analyzed neighborhood for every landscape type and to choose the highest one to the k-th landscape type.

The local characteristics of every k-th landscape type are calculated using the 7 mentioned characteristics in the analyzed neighborhood and can be presented as a vector

$$\mathbf{F}_{\mathrm{C\,loc}} = [C^k_{\mathrm{FO}}\, C^k_{\mathrm{LF}}\, C^k_{\mathrm{AFM}}\, C^k_{\mathrm{ADP}}\, C^k_{\mathrm{LD}}\, C^k_{\mathrm{LPS}}\, C^k_{\mathrm{AFSPA}}]^{\mathrm{T}}.$$

The global characteristics can also be presented as a vector

$$\mathbf{F}_{\mathbf{Cglb}} = [\,\overline{C}^{\,k}_{\mathrm{FO}}\ \overline{C}^{\,k}_{\mathrm{LF}}\ \overline{C}^{\,k}_{\mathrm{LF}}\ \overline{C}^{\,k}_{\mathrm{AFM}}\ \overline{C}^{\,k}_{\mathrm{ADP}}\ \overline{C}^{\,k}_{\mathrm{LD}}\ \overline{C}^{\,k}_{\mathrm{LPS}}\ \overline{C}^{\,k}_{\mathrm{AFSPA}}\ ]^{T}.$$

The components of the vector $\mathbf{F}_{\mathbf{Cglb}}$ are calculated for every k-th type of the landscape by moving the sliding window throughout the image and calculating the average value of every characteristic for the whole image. The proposed way of defining CA transition rules on the basis of p_{ij}^{res} expression and p^{sp} expression was included in an algorithm of land cover change modeling implemented in the sophisticated software of land covers change analysis.

3 Results and Discussion

To determine the accuracy performance of the land cover change modeling algorithm hundreds of experiments were performed. The experiments were directed to search for spatial characteristics informativeness in varying neighborhood size and their significance level in different combinations. The time series *Idrisi Kilimanjaro* sample thematic maps were used: the modeling itself performed on 1971 and 1985 maps; the ground truth data – 1991 map; *Receiver/Relative Operating Characteristic* (ROC) and *Kappa Index of Agreement* (KIA) accuracy criteria were used (RICHARDS & XIUPING 1999; WRIGHT 2005). ROC curve is a graphical plot of the number of true positives versus the number of false positives for a binary classifier system as its discrimination threshold is varied, and it is often used to evaluate the results of a prediction. KIA is a measure that is mainly used in the field of remote sensing to compare two images with a contingency table to find whether their differences are due to chance or real disagreement or agreement.

To define spatial characteristics informativeness a set of experiments with 3×3, 5×5, 7×7, 9×9, 15×15 and 29×29 elements in the neighborhood were performed. The modeling was performed with 7 characteristics separately, as well as without any spatial characteristic at all. The experiments allow not only to prove the appropriateness of spatial characteristics use in CA modeling (the accuracy of modeling results with any characteristic is significantly higher than the same results obtained without including the characteristic), but to find out *sensibility* (level of significance) of every single characteristic associated with each different neighborhood size. The research demonstrates that nearly the best accuracy results in all cases were found for the window size with 5×5 elements for both KIA and ROC accuracy criteria. It is obvious, that the impact of spatial characteristics may significantly depends on both the neighborhood size chosen and the features (e.g. the scale) of specific landscape patterns. The estimation (particularly visual) of obtained results for different neighborhood size confirmed that too small neighborhood size as 3×3 may allow taking into account landscape spatial features to some extent only. But if we considerably increase the neighborhood size, the generalization level of spatial features becomes too high and spatial features do not distinguish different points of a landscape sharply. Hence, as a compromise for the ability to distinguish landscapes features it is proposed to use the 5×5 neighborhood size for similar land cover change modeling tasks.

Exceptions were the results obtained with C_{AFM} and C_{LPS} characteristics that did not show robust correlation with neighborhood size. Besides the numerical method of informativeness evaluation, the visual method of the evaluation was used. The visual analysis of images demonstrates the fact that every characteristic allows to reveal some unique spatial features of the analyzed image and to point out the borders between landscape types.

Every spatial characteristic allows revealing unique features of the analyzed landscape due to which their practical use is reasonable to perform together, taking into account the set of different spatial features in a neighborhood. The author proposes to search for the most effective set of spatial characteristics in the following way.

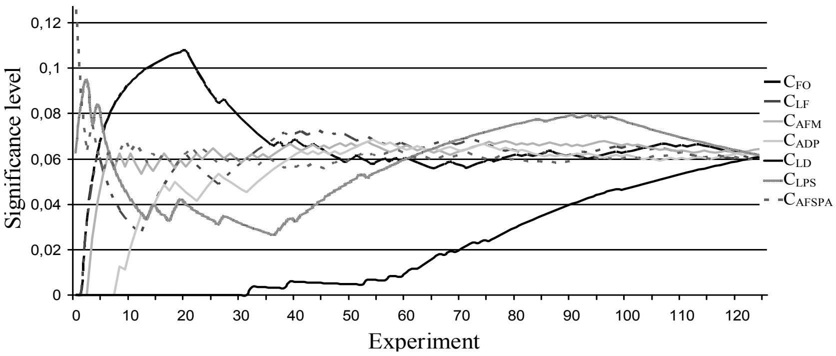

Fig. 1: Significance level of every characteristic for different combinations of spatial characteristics ranked in descending order according to KIA accuracy criterion value.

The overall number of experiments for 7 characteristics evaluation process is $N_{expr}=2^7=128$ where each experiment is performed using a unique combination of characteristics. A result of every experiment is a forecast map. Comparing it with the ground truth map we may obtain the evaluation results according to the ROC and KIA criteria. After that, the results should be arranged in descending order according to accuracy criterion value (in this case the KIA criterion is chosen). A significance level of C characteristic in a single k experiment is proposed to evaluate as $S_C^k = v_C^k \cdot acc^k / M$, where k=1,2,…,N_{expr}, acc – value of accuracy criterion (in this case the KIA value), M – the number of characteristics used in modeling (in this case M=7), v={0,1}: v=0, if C is absent in the set of characteristics combination and v=1 – if C is present. The significance level of C according to the results of N experiments (integral evaluation) can be defined as $S_C^N = 1/N \sum_{k=1}^{N} S_C^k$. Figure 1 shows results of significance level evaluation for each characteristic where N=1,2,…,N_{expr}. The analysis of dependences presented in Fig. 1 shows that the most precise results of modeling are obtained using combinations of characteristics C_{FO}, C_{LF}, C_{LPS} and C_{AFSPA}. It demonstrates the higher significance level of these characteristics in conducted experiments. The smaller significance level of C_{AFM}, and C_{ADP}, C_{LD} characteristics is proved by numerical and visual evaluation.

4 Conclusions

The results of the conducted research confirmed preliminary conclusions obtained by the author earlier, and showed the appropriateness of using spatial characteristics to define cellular automata transition rules in order to increase the accuracy of land cover change modeling. The experiment results for informativeness evaluation of 7 spatial characteristics in a different neighborhood size, obtained with the time series data samples of *Idrisi Kilimanjaro*, show that definition of cellular automata transition rules is more reasonable to perform due to practical reason in the neighborhood size with 5×5 elements. The method is proposed for significance level definition of spatial characteristics in different sets throughout cellular automata land cover change modeling. The method allows to underline the characteristics frequency of occurrence (C_{FO}), level of fragmentation (C_{LF}), level of patch size (C_{LPS}) and average fractal size of perimeters-areas (C_{AFSPA}) having the highest significance level according to accuracy criterion values.

References

BENENSON, I. & TORRENS, P. (2004), Geosimulation: Automata-based modeling of urban phenomena. Chichester, Wiley.

LI, X. & GAR-ON YEH, A. (2004), Data mining of cellular automata's transition rules. Int. J. Geographical Information Science, 18 (8), pp. 723-744.

RICHARDS, J. A. & XIUPING, J. (1999), Remote Sensing Digital Image Analysis: An Introduction. Berlin, Springer.

TISCHENDORF, L. (2001), Can landscape indices predict ecological processes consistently? Landscape Ecology, 16 (3), pp. 235-254.

VERBURG, P. H. et al. (2003), A method to analyse neighborhood characteristics of land use patterns. Computers, Environment and Urban Systems, 24, pp. 354-369.

WRIGHT, D. B. (2005), Receiver Operating Characteristics Curves. Encyclopedia of Statistics in Behavioral Science. Chichester, Wiley, 4, pp. 1718-1721.

ZAMYATIN, A., MIKHAYLOV, P. & CABRAL, P. (2007), Spatial Metrics Approach to Land Cover Change Forecasting Using Cellular Automata. AGILE 2007 Conf. Proc. Available: http://people.plan.aau.dk/~enc/AGILE2007/PDF/90_PDF.pdf (01.03.2009).