HPC-ICTM: the interval categorizer tesselation-based model for high perfomance computing*

Marilton S. de Aguiar[1], Graçaliz P. Dimuro[1], Antônio C. da R. Costa[1]
Rafael K. S. Silva[1] and Vladik Kreinovich[2]

[1] Escola de Informática
Universidade Católica de Pelotas
Felix da Cunha, 412. CEP 96010-000. Pelotas/RS, Brazil
{marilton,liz,rocha,rafaelk}@ucpel.tche.br
[2] Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
vladik@cs.utep.edu

Abstract. The paper presents the results obtained by an implementation of the interval tessellation-based model for categorization of geographic regions according the analysis of the relief function declivity, called ICTM. The analysis of the relief declivity, which is embedded in the rules of the model ICTM, categorizes each tessellation cell, with respect to the whole considered region, according to the (positive, negative, null) signal of the declivity of the cell. Such information is represented in the states assumed by the cells of the model. The overall configuration of such cells allows the division of the region into sub-regions of cells belonging to the same category, that is, presenting the same declivity signal. In order to control the errors coming from the discretization of the region into tessellation cells, or resulting from numerical computations, interval techniques are used.

1 Introduction

The tessellation-based model performs a bi-dimensional analysis of the declivity, using local rules for creation and categorization of sub-regions, giving the relative situation of each sub-region with respect the whole area, according to the

* This work was supported by the Brazilian funding agencies CNPq/CTPETRO/CTINFO and FAPERGS. V.K. was also partly supported by NASA under cooperative agreement NCC5-209, by Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant F49620-00-1-0365, by NSF grants EAR-0112968 and EAR-0225670, by the Army Research Laboratories grant DATM-05-02-C-0046, and by IEEE/ACM SC2003 Minority Serving Institutions Participation Grant.

states assumed by the cells. This work evolved directly from the analysis of the work [3]. The `ICTM` Model uses a structured mesh to constitute its tesselation. A structured bi-dimensional mesh is often simply a square grid deformed by some coordinate transformation. Each vertex of the mesh, except at the boundaries, has an isomorphic local neighborhood. In three dimensions, a structured mesh is usually a cubical grid. Structured meshes are simpler than the non-structured ones, and require less computer memory, as their coordinates can be calculated, rather than explicitly stored. Structured meshes offer more direct control over the sizes and shapes of elements.

An immediate application is in Geophysics, where an adequate subdivision of geographic areas into segments presenting similar topographic characteristics is often convenient. See [5], for other applications related to the analysis of the relief. The data input for the model are extracted from satellite images of the geographic region being analyzed, where the heights are given in certain points referenced by their latitude and longitude coordinates. This geographic region is represented by a regular tessellation that is determined by subdividing the total area into sufficiently small rectangular subareas, each one represented by one cell of the tessellation. This subdivision is done according to a cell size established by the geophysics analyst and it is directly associated to the refinement degree of the tessellation.

The categorization determined by each characteristic is performed in one layer of the model, generating different subdivisions of the analyzed region. For instance, a region can be analyzed according to its topography, vegetation, demography, economic data, etc. A global categorization can be reached from the categorization of each layer through a projection procedure. This global categorization will determinate a more reliable and significant subdivision combining the performed analysis in each characteristic.

2 The Formalization of the `ICTM` Model

This section introduces the multi-layered interval categorizer tessellation-based model, formalized in terms of matrix operations. The `ICTM` single-layered was firstly presented in [1]. Here, we present the generalization on the number of the layers and the projection procedures. This type of projection allows interesting analysis about the mutual dependency of this characteristics. Each characteristic of the space is represented in a layer of the `ICTM` Model. Thus, by the independency of the analysis, the subdivisions in each layer also occurs in a independently way.

Definition 1 *A tessellation is a matrix M with n_r rows and n_c columns. The entry at the x-th row and the y-th column is called the xy-cell of M.*

Definition 2 *Considering a $n_c \times n_r$ tesselation M and $l \in \mathbb{N}$, a multi-layered tesselation $\mathcal{L}$-$\mathcal{M}$ is the structure*

$$\mathcal{L}\text{-}\mathcal{M} = \left(1\text{-}M^{abs}, \ldots, l\text{-}M^{abs}\right)$$

where the entry at the l-th layer, x-th row and y-th column is denoted by l-m_{xy}.

2.1 The Interval Matrices

In topographic analysis, usually there are too much data, most of which is geophysically irrelevant. We then take, for each subdivision, the average value of the heights at the points, which are the entries of the tessellation M:

Definition 3 *A layer l of this tesselation M is the $n_r \times n_c$ matrix $l\text{-}M^{abs} = [l\text{-}m_{xy}^{abs}]$, where the entry $l\text{-}m_{xy}^{abs}$ is the absolute value of the average height of the points represented by the xy-cell in the layer l of M.*

To simplify the data of the matrix, we normalize them by dividing each $l\text{-}m_{xy}^{abs}$ by the largest $l\text{-}m_{max}$ of these values.

Definition 4 *The relative matrix of layer l $l\text{-}M^{rel}$ is defined as the $n_r \times n_c$ matrix given by*

$$l\text{-}M^{rel} = \frac{l\text{-}M^{abs}}{l\text{-}m_{max}}.$$

In the following, we apply Interval Mathematics [6] techniques to control the errors associated to the cell values[3]. For each $\xi\upsilon$, which is different from xy, it is reasonable to estimate $h_{\xi\upsilon}$ as the value $l\text{-}m_{xy}^{rel.}$ at the point xy which is closest to $\xi\upsilon$, meaning that $\xi\upsilon$ belongs to the same segment of area as xy. For each cell xy, let Δ_x and Δ_y be the largest possible errors of the corresponding approximations considering the west-east direction and the north-south direction, respectively.

Lemma 1 *For fixed y, if $\xi > x$, then the approximation error ϵ is bounded by* $0.50 \cdot |l\text{-}m_{(x+1)y}^{rel} - l\text{-}m_{xy}^{rel}|$.

Lemma 2 *For fixed y, if $\xi < x$, then the approximation error ϵ is bounded by* $0.50 \cdot |l\text{-}m_{xy}^{rel} - l\text{-}m_{(x-1)y}^{rel}|$.

Proposition 1 *For the approximation error ϵ_x,*

$$\epsilon_x \leq \Delta_x = 0.5 \cdot \min\left(|l\text{-}m_{xy}^{rel} - l\text{-}m_{(x-1)y}^{rel}|, |l\text{-}m_{(x+1)y}^{rel} - l\text{-}m_{xy}^{rel}|\right).$$

Proof. It follows from Lemmas 1 and 2.□

As a result, considering a given y, besides of the central values $l\text{-}m_{xy}^{rel}$, for each x, we get intervals $m_{xy}^{x[\,]}$ containing all the possible values of $h_{\xi y}$, for $x - \frac{1}{2} \leq \xi \leq x + \frac{1}{2}$.

Corollary 1 *Considering a fixed y, for each x, if $x - \frac{1}{2} \leq \xi \leq x + \frac{1}{2}$, then $h_{\xi y} \in l\text{-}m_{xy}^{x[\,]} = \left[l\text{-}m_{xy}^{x^-}, l\text{-}m_{xy}^{x^+}\right]$, where $l\text{-}m_{xy}^{x^-} = l\text{-}m_{xy}^{rel} - \Delta_x$ and $l\text{-}m_{xy}^{x^+} = l\text{-}m_{xy}^{rel} + \Delta_x$.*

Using an analogous argumentation, for a fixed x, it follows that:

[3] See examples of using intervals in solving similar problems in [3, 4].

Proposition 2 *For the approximation error ϵ_y,*

$$\epsilon_y \leq \varDelta_y = 0.5 \cdot \min\left(|l\text{-}m^{rel}_{xy} - l\text{-}m^{rel}_{x(y-1)}|, |l\text{-}m^{rel}_{x(y+1)} - l\text{-}m^{rel}_{xy}|\right).$$

Corollary 2 *Considering a fixed x, for each y, if $y - \frac{1}{2} \leq \upsilon \leq y + \frac{1}{2}$, $h_{x\upsilon} \in l\text{-}m^{y^{[\,]}}_{xy} = \left[l\text{-}m^{y^-}_{xy}, l\text{-}m^{y^+}_{xw}\right]$, where $l\text{-}m^{y^-}_{xy} = l\text{-}m^{rel}_{xy} - \varDelta_y$, $l\text{-}m^{y^+}_{xy} = l\text{-}m^{rel}_{xy} + \varDelta_y$.*

Definition 5 *If $l\text{-}m^{x^\pm}_{xy} = l\text{-}m^{rel}_{xy} \pm \varDelta_i$ and $l\text{-}m^{y^\pm}_{xy} = l\text{-}m^{rel}_{xy} \pm \varDelta_j$, the interval matrices $l\text{-}M^{x^{[\,]}}$ and $l\text{-}M^{y^{[\,]}}$, associated with the relative matrix $l\text{-}M^{rel}$, are defined by the $n_r \times n_c$ interval matrices*

$$l\text{-}M^{x^{[\,]}} = \left[l\text{-}m^{x^{[\,]}}_{xy}\right] = \left[\left[l\text{-}m^{x^-}_{xy}, l\text{-}m^{x^+}_{xy}\right]\right], l\text{-}M^{y^{[\,]}} = \left[l\text{-}m^{y^{[\,]}}_{xy}\right] = \left[\left[l\text{-}m^{y^-}_{xy}, l\text{-}m^{y^+}_{xy}\right]\right].$$

2.2 The Declivity Registers and the State Matrix

We proceed to a declivity categorization[4] assuming that the relief approximation functions introduced by the tessellation-based model are piecewise linear functions. We cast the whole process as a kind of constraint satisfaction problem, where the tessellation-based model is in charge of finding a piecewise linear relief approximation function (and corresponding set of limit points between the resulting sub-regions) that fits the constraints imposed by the interval matrix. To narrow the solution space to a minimum, we take a qualitative approach to the relief approximation functions, clustering them in equivalence classes according to the signal of their declivity (positive, negative, null), thus making the tessellation-based model build a single qualitative solution to that constraint satisfaction problem, namely, the class of approximation functions compatible with the constraints of the interval matrix. We proceed as follows:

Proposition 3 *Let $l\text{-}M^{x^{[\,]}}$ and $l\text{-}M^{y^{[\,]}}$ be interval matrices of layer l. For a given xy, if:*

(i) $l\text{-}m^{x^+}_{xy} \geq l\text{-}m^{x^-}_{(x+1)y}$, *then there exists a non-increasing relief approximation function between xy and $(x+1)y$ (direction west-east).*
(ii) $l\text{-}m^{x^-}_{(x-1)y} \leq l\text{-}m^{x^+}_{xy}$, *then there exist a non-decreasing relief approximation function between $(x-1)y$ and xy (direction west-east).*
(iii) $l\text{-}m^{y^+}_{xy} \geq l\text{-}m^{y^-}_{x(y+1)}$, *then there exists a non-increasing relief approximation function between xy and $x(y+1)$ (direction north-south).*
(iv) $l\text{-}m^{y^-}_{x(y-1)} \leq m^{y^+}_{xy}$, *then there exists a non-decreasing relief approximation function between $x(y-1)$ and xy (direction north-south).*

Definition 6 *A declivity register of an xy-cell is a tuple*

$$reg = (reg.e, reg.w, reg.s, reg.n)$$

where the values of the directed declivity registers are given by:

[4] This declivity categorization was inspired by [3].

(a) For non border cells, considering the conditions given by Prop. 3: $reg.e = 0$, if (i) holds; $reg.w = 0$, if (ii) holds; $reg.s = 0$, if (iii) holds; $reg.n = 0$, if (iv) holds; $reg.e, reg.w, reg.s, reg.n = 1$, otherwise.
(b) For east, west, south and north border cells: $reg.e = 0$, $reg.w = 0$, $reg.s = 0$ and $reg.n = 0$, respectively[5]. The other directed declivity registers of border cells are also determined according to item (a).

Definition 7 *The declivity register matrix of the layer l is defined as an $n_r \times n_c$ matrix $l\text{-}M^{reg} = \left[l\text{-}m_{xy}^{reg}\right]$, where the entry at the x-th row and the y-th column is the value of the declivity register of the corresponding cell.*

Corollary 3 *Considering the west-east direction, any relief approximation function $l\text{-}m_{xy}$ is either (i) strictly increasing between xy and $(x+1)y$ if $l\text{-}m_{xy}^{reg.e} = 1$ (in this case, $l\text{-}m_{(x+1)y}^{reg.w} = 0$); or (ii) strictly decreasing between xy and $(x+1)y$ if $l\text{-}m_{(x+1)y}^{reg.w} = 1$ (in this case, $l\text{-}m_{xy}^{reg.e} = 0$); or (iii) constant between xy and $(x+1)y$ if $l\text{-}m_{xy}^{reg.e} = 0$ and $l\text{-}m_{(x+1)y}^{reg.w} = 0$. Similar results hold for the north-south direction.*

Definition 8 *Let $w_{reg.e} = 1$, $w_{reg.s} = 2$, $w_{reg.w} = 4$ and $w_{reg.n} = 8$ be weights to be associated to the directed declivity registers. The state matrix is defined as an $n_r \times n_c$ matrix given by $l\text{-}M^{state} = \left[l\text{-}m_{xy}^{state}\right]$, where the entry at the x-th row and the y-th column is the value of the corresponding cell state, calculated as the value of the binary encoding of the corresponding directed declivity registers, given as*

$$l\text{-}m_{xy}^{state} = w_{reg.e} \times l\text{-}m_{xy}^{reg.e} + w_{reg.s} \times l\text{-}m_{xy}^{reg.s} + w_{reg.w} \times l\text{-}m_{xy}^{reg.w} + w_{reg.n} \times l\text{-}m_{xy}^{reg.n}.$$

Thus, for given xy, the correspondent cell can assume one and only one state represented by the value $l\text{-}m_{xy}^{state} = 0..15$.

2.3 The Limiting Matrix and the Constant-Declivity Sub-Regions

A limiting cell is defined as the one where the relief function changes its declivity, presenting critical points (maximum, minimum or inflection points). To identify such limiting cells, we use a limiting register associated to each cell. The border cells are assumed to be limiting.

Definition 9 *The limiting matrix of the layer l is defined as the $n_r \times n_c$ matrix given by $l\text{-}M^{limit} = \left[l\text{-}m_{xy}^{limit}\right]$, where the entry at the x-th row and the y-th column is determined as $l\text{-}m_{xy}^{limit} = 0$, if one of the conditions listed in Table 1 holds, and $l\text{-}m_{xy}^{limit} = 1$, otherwise.*

Definition 10 *The constant declivity sub-region associated to the non limiting cell xy, denoted $l\text{-}\mathcal{SR}_{xy}$, is inductively defined as follows:* (i) *$xy \in l\text{-}\mathcal{SR}_{xy}$;* (ii) *If $x'y' \in l\text{-}\mathcal{SR}_{xy}$, then all its neighbor cells that are not limiting cells also belong to $l\text{-}\mathcal{SR}_{xy}$.*

[5] This is consistent with the relief function being a constant in the border cells.

Table 1. Conditions of non limiting cells

Id	Conditions
1	$l\text{-}m^{reg.e}_{(x-1)y} = l\text{-}m^{reg.e}_{xy} = 1$
2	$l\text{-}m^{reg.w}_{xy} = l\text{-}m^{reg.w}_{(x+1)y} = 1$
3	$l\text{-}m^{reg.e}_{(x-1)y} = l\text{-}m^{reg.e}_{xy} = l\text{-}m^{reg.w}_{xy} = l\text{-}m^{reg.w}_{(x+1)y} = 0$
4	$l\text{-}m^{reg.s}_{x(y-1)} = l\text{-}m^{reg.s}_{xy} = 1$
5	$l\text{-}m^{reg.n}_{xy} = l\text{-}m^{reg.n}_{x(y+1)} = 1$
6	$l\text{-}m^{reg.s}_{x(y-1)} = l\text{-}m^{reg.s}_{xy} = l\text{-}m^{reg.n}_{xy} = l\text{-}m^{reg.n}_{x(y+1)} = 0$

Table 2. Number of Categories

Radius	DEM	
	1000m	500m
1	76	230
2	62	197
5	36	143
10	22	125
20	18	108
40	18	83

Observe that $l\text{-}\mathcal{SR}_{xy} = l\text{-}\mathcal{SR}_{x'y'}$ if and only if $x'y' \in l\text{-}\mathcal{SR}_{xy}$ (resp., $xy \in l\text{-}\mathcal{SR}_{x'y'}$). Definition 10 leads to a recursive algorithm similar to the ones commonly used to fulfill polygons.

3 Some practical results

This section presents some results[6] that this work already reached and also some notes for future works. Up the moment, the model of parallel processing that is being implemented creates independent processes for each analyzed property, performing sequentially in each one of these processes all the `ICTM` activities (presented in the sections 2.1– 2.3). However, is also possible to wait for a performance enhancement when the proper activities of the `ICTM` model will be processed in parallel. The tessellation can be divided to be processed separately using an algorithm (i.e. the Schwarz algorithm [2]) for overlapping domain decomposition. Domain decomposition methods are techniques for solving partial differential equations based on a decomposition of the spatial domain of the problem into several subdomains. The implementation of the model is naturally parallel since the analysis is performed on the basis of local rules. Our implementation, taking advantage of such characteristic, uses the MPI standard on top of a distributed processing cluster. A discussion of the performance of the parallel model and a comparison with the sequential implementation will be presented in the future.

The results below use Digital Elevation Models (DEM) of resolutions 1000 and 500 meters, with the following coordinates[7]: (i) Upper-left corner at ($X = 427559m$, $Y = 6637852m$) and (ii) Lower-right corner at ($X = 480339m$, $Y = 6614507m$). It can be observed (see Table 2 and Fig. 1) that the number of categories is inversely proportional to the neighborhood radius. Moreover, for the 500m resolution, where each point has a smaller area (approx. four times) than the corresponding point in DEM of a 1000m resolution, the number of categories also follow (approx.) this factor. In this case, the DEM of 500m resolution (of

[6] The results can be viewed in more detail at http://descartes.ucpel.tche.br/ictm.

[7] These coordinates are UTM 22S (South Hemisphere) and Datum SAD69 (South America Datum)

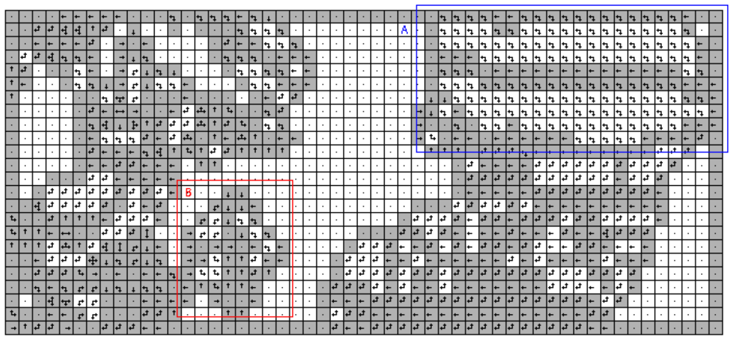

Fig. 1. In plain areas (region A), bigger neighborhood radius indicate reasonable approximations for this declivity degree. However, regions with bigger declivity (region B) variations obtained better aproximations with smaller radius. In regions with bigger number of categories are indicative that an analysis more detailed must be done.

this region) does not present new sub-areas to be categorized, just showing the sub-areas in more detail. However, if the region has much declivity variation then probably this factor shall not be verified anymore. The ICTM Model is regulated by two aspects: (i) the spacial resolution of the digital elevation model, and (ii) the neighborhood radius of the cell. Thus, regions with an agglomeration of limiting cells can be studied with more details by just increasing the resolution of altimetry data, or reducing the neighborhood radius. In the ICTM model, the state of a cell in relation to its neighbors, concerning the declivity, can be verified instantaneously, contrasting with the usual analysis. The results indicate that regions with less declivity variation are receptive to the ICTM model with bigger radius. In contrast, regions with more declivity variation suggest smaller radius.

References

1. M. S. Aguiar, A. C. R. Costa, and G. P. Dimuro. ICTM: an interval tesselation-based model for reliable topographic segmentation. *Numerical Algorithms*, 2004. (to be published).
2. Tony F. Chan and Tarek P. Mathew. Domain decomposition algorithms. In *Acta Numerica*, pages 61–143. Cambridge University Press, 1994.
3. D. Coblentz, V. Kreinovich, B. Penn, and S. Starks. Towards reliable sub-division of geological areas: Interval approach. In L. Reznik and V. Kreinovich, editors, *Soft Computing in Measurements and Information Acquisition*, pages 223–233. Springer-Verlag, Berlin-Heidelberg, 2003.
4. R. B. Kearfort and V. Kreinovich (eds.), *Aplications of Interval Computations*, Kluwer, Dordrecht, 1996.
5. R. T. T. Forman. *Land Mosaics: the ecology of landscapes and regions.* Cambr. Univ. Press, Cambridge, 1995.
6. R. E. Moore. *Methods and Applications of Interval Analysis.* SIAM, Philadelphia, PA, USA, 1979.