

# On the Generation of Aggregated Random Spatial Regions \*

Yannis Manolopoulos<sup>1</sup>

Enrico Nardelli<sup>2,3 †</sup>

Guido Proietti<sup>2,3 †</sup>

Michael Vassilakopoulos<sup>4</sup>

1. Department of Informatics, Aristotle University of Thessaloniki, GR-54006, Thessaloniki, Greece, e-mail: manolopo@eng.auth.gr.
2. Dipartimento di Matematica Pura ed Applicata, Università di L'Aquila, Via Vetoio, Loc. Coppito, I-67010 L'Aquila, Italia.
3. Istituto di Analisi dei Sistemi ed Informatica, Consiglio Nazionale delle Ricerche, Viale Manzoni 30, I-00185 Roma, Italia, e-mail: {nardelli, guido}@iasi.rm.cnr.it.
4. Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, GR-54006 Thessaloniki, Greece, e-mail: vasilako@eng.auth.gr.

## Abstract

Traditional random models for spatial two-dimensional data proposed in literature show their limits in generating in a satisfactory way instances of regions having a desired aggregation level. This is because none of them is really oriented to this aim. Rather, they are thought to model the behaviour of the constituting elements of the spatial data (so losing sight of the context), or, alternatively, to model particular data structure for their representation, underestimating the fact that there is in general no semantic link between a region data and its representation. This means from one hand, the impossibility to produce meaningful theoretical results on time and space average performances of different data structures used to represent spatial regions, and, on the other hand, in an applicative context, the difficulty to generate instances of spatial regions having a statistical behaviour close to that of real data. To overcome this trouble, we introduce in our paper a new random model that provides the possibility to generate spatial regions having a desired aggregation.

## 1 Introduction

Two dimensional binary region data (simply *spatial regions* in the following) can be considered as an  $n \times n$  array of constituting elements (termed *points* or *cells*), each of which belongs or not to the region itself. Many different approaches have been proposed in the literature to represent regions: array representation (raster-based), run-length codes, polygons (vector-based), bounding boxes, mapping to higher or lower dimensional spaces, region quadtrees and so on; an interested reader may refer to [6] for a survey.

\*Research performed under the 4th Italian-Greek bilateral scientific protocol.

<sup>†</sup>Partially supported by the Italian MURST 40% Project "Algoritmi, Modelli di Calcolo e Strutture Informative" and by the Italian CNR Coordinated Project "Modelli e Sistemi per il Trattamento di Dati Ambientali e Territoriali".

Permission to make digital/hard copies of all or part of this material for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copyright is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires specific permission and/or fee.

CIKM '95, Baltimore MD USA

© 1995 ACM 0-89791-812-6/95/11..\$3.50

Depending from the applicative context, one representation is preferred to the others. This is because the "morphology" of the spatial regions is strongly tied to the application: so, for example, landuse maps generally store pseudo-rectangular regions, while topographic maps are a mixture of geometric shapes and lines, since they have to represent buildings and networks as the electric or the drainage nets; on the contrary, pebble maps appear the most of times as the spots of a leopard. Then, the choice of the optimal representation (in the sense of the best compromise between time and space efficiency) requires a preliminary work in which we test the efficiency of the various approaches on a large number of cases. From a different perspective, the introduction of a new data representation should be justified exploiting its qualities with respect to the state of the art and the class of data under consideration, and this can be made properly only performing statistics on a big enough number of cases. This means the need to have at disposal a random generator of spatial regions, good enough to produce instances with the morphological peculiarities of the considered class of data. In fact, this avoids the irksome task of finding suitable data among the large volumes of spatial data having different and often unknown nature which can be obtained from sites scattered around the world. We have then to overcome the not easy problem to give a random model describing in a satisfactory way the intrinsic "structure" of a spatial region. This has led in the literature to the definition of a number of random models aiming to perform such a task [3, 4, 5, 7, 8].

The first and most famous random model for regions is the conventional *pixel based* model [4, 7], so called since it is designed to describe regions in terms of representing images, given the large use that is made of pictures to store two-dimensional spatial data set. In this model, each point of the region is associated to a pixel in the image and each of these pixels is assumed to be statistically independent from any other pixel; if  $\rho$  is the probability for a generic point to belong to the region and then for a pixel to be black, then a level- $i$  quadrant (i.e., a quadrant of width  $2^i$ ) is completely black with probability  $B_i = \rho^{4^i}$ , completely white with probability  $W_i = (1 - \rho)^{4^i}$ , while it is gray with probability  $G_i = 1 - \rho^{4^i} - (1 - \rho)^{4^i}$ . This model leads to a very low probability of *aggregation*. With the term *aggregation* we informally mean an overall measure of the average mutual adjacency among the black points composing the

spatial region. The problem with the pixel based model is that it assumes independence among pixels, while the real case in spatial data sets is that a point's value is typically related to that of its neighbours. So, even if the probability for a pixel to be black is high, the image tends always to be sparse and consequently to be far from typical instances of most of classes of spatial regions (e.g., landuse maps, geographical maps and also sets of geometrical objects). This implies that, in spite of its usefulness in a image processing context, the pixel based model is from a practical point of view useless to work on spatial regions.

A very popular alternative approach that overcomes this disadvantage is the so called *tree based* model [4, 5], that aims to model the quadtree representing the spatial data set. In this model, given a non increasing sequence  $\beta = (\beta_0, \dots, \beta_m)$  of  $m + 1$  reals such that  $0 < \beta_i < 1/2$  but  $\beta_0 = 1/2$ , a random quadtree  $Q_m$  of height  $m$  is built by using a branching process, such that the quantity  $2\beta_i$  is the probability for a level  $i$  node to be a leaf (and so by definition  $2\beta_0 = 1$ , i.e., at level 0 all nodes are leaves with probability 1 as they should), and then  $G_i = 1 - 2\beta_i$  is the probability for a level  $i$  node to be internal. Once a node is known to be a leaf, it can indifferently contain or not the spatial region.

This tree based model allows to generate very "compact" quadtrees, and then regions having a high degree of aggregation, also in spite of a small regions area with respect to the background. Nevertheless, this approach has a big deficiency too, residing in the fact that it models the data structure and not the spatial region! Even if it is generally true that dense quadtrees are associated to aggregated regions, it is not true the viceversa, that is very aggregated regions could be represented (depending by their spatial position) by very "bushy" quadtrees. This happens because the same spatial data set moved inside the image space can generate quadtrees having completely different structures [1]. In [2] exact formulae are derived for the estimation of the average number of quadtrees nodes required by a rectangle aligned with the axes. So, if we are interested in studying a particular class of spatial data sets that we know in advance to have a high degree of aggregation, it will not be sufficient to define a branching sequence aiming to generate compact quadtrees, since a large class of instances will be at last cut off. Figure 1 shows how the same region moved inside the image space can generate completely different quadtrees.

A further problem with the tree based model is that it creates non-condensed quadtrees (a node may have four children that are all white or all black). This specific problem has been overcome in [8], where an extension of the tree based model that only creates condensed quadtrees is given. To solve on the contrary the problem of linking the building of the quadtree to the desired shape of the region, we should mould the branching process on the basis of a probability function describing the distribution of the region in the image space. A first attempt in this sense has been made in [3], but the process there suggested can be efficiently applied only when the morphology of the region can be described by a probability function, and this is not a very frequent case.

Then, current random models for spatial data cannot be efficiently applied to work on spatial regions, since they do not adequately represent, as previous examples shows, the investigated class of data. Hence, it is clear that a novel approach is needed. We then provide a new random model able to take directly into account the degree of aggregation inside a region. Starting from the traditional pixel independent model, we mould the generated set of points until a desired level of aggregation is reached. We show how the

flexibility of this new model is useful in an applicative context, especially whenever it is needed to focus on subclasses of homogeneous spatial data sets, as for example landuse or topographical maps.

The paper proceeds as follows: in section 2 we give necessary definitions and introduce the new model; in section 3 we provide the algorithm to generate random regions using the new approach; in section 4 we show how much flexible the new model is, presenting experimental results; finally section 5 contains conclusions and open problems.

## 2. The new model

By using the term *aggregation* of spatial regions, we refer to an overall measure deriving from the adjacency among the black cells composing the spatial region. The adjacency between two black cells can be *isothetic*, if it happens in the horizontal or vertical direction, or it can be *diagonal*, if the two cells touch each other in a corner. In what follows we will not distinguish between the two kinds of adjacency, and we will say that two cells are *adjacent* if they are isotetically or diagonally adjacent, while we will call *disjoint* two cells that are not adjacent. A black cell will be called *single-point* if it does not have any adjacent black cell.

The question is now the following:

Given an integer  $k$ , which is the spatial region  $\mathcal{I}_k$  made up of  $k$  black cells to consider as the one having the maximal degree of aggregation?

To answer to this question, we have to introduce some definitions. Given a black cell  $p$  belonging to  $\mathcal{I}_k$ , we associate to it an *adjacency number*  $adj(p) = b/8$ , where  $b$  is the number of black cells adjacent to  $p$ . This value is minimum (i.e.,  $adj(p) = 0$ ) when  $p$  is a single-point, while it is maximum (i.e.,  $adj(p) = 1$ ) when  $p$  is surrounded by only black cells. It is obvious that the greater is the sum over all the black cells of the respective adjacency number, the more the spatial region is aggregated. We therefore introduce the *aggregation factor*:

$$agg(\mathcal{I}_k) = \sum_{p \in \mathcal{I}_k} adj(p)$$

The problem is that this number depends on the number of black cells constituting the spatial region itself. To overcome this problem, we normalize the value with respect to the maximal obtainable value of  $agg(\mathcal{I}_k)$  for every spatial region  $\mathcal{I}_k$  that is made up of  $k$  black cells. Once we individuate such a spatial region, that we indicate with  $\mathcal{I}_k^*$ , we will define the *coefficient of aggregation*  $\alpha(\mathcal{I}_k)$  of a spatial region  $\mathcal{I}_k$  made up of exactly  $k$  cells as:

$$\alpha(\mathcal{I}_k) = \frac{agg(\mathcal{I}_k)}{agg(\mathcal{I}_k^*)}$$

This number expresses the normalized degree of adjacency on each black cell belonging to the spatial region; it is clearly  $0 \leq \alpha(\mathcal{I}_k) \leq 1$ .

To find the normalization factor  $agg(\mathcal{I}_k^*)$  we could proceed by generating all spatial regions made up of  $k$  black cells and calculating for each one of them the respective aggregation factor, but of course this is not practical, given that the number of such spatial regions is exponential in  $k$ .

We start then by observing that the spatial region having maximal aggregation factor has to be 8-connected. In fact, it is obvious that if the spatial region is composed by,

for example, two disconnected components, then the aggregation factor grows joining regions, since at least one cell in each region will have its adjacency number increased and no cell in any region will have its adjacency number decreased.

Aftewards, to obtain a more precise indication on the value of  $agg(\mathcal{I}_k^*)$  we follow the approach used in the euclidean plane to solve the minimal perimeter problem (having in mind that this is not a formal correspondence, given the dramatic differences between the two frameworks). In fact, in the euclidean plane one wants to find, among all figures having the same area, the one which maximizes adjacency among points. This is the one having minimum perimeter, since points on the boundary are adjacent to external points. To find the minimum perimeter figure, an argument based on symmetry can be used. Divide the figure in two parts with an axis such that the areas of the two parts are equal. Take the part having a larger perimeter (or choose arbitrarily one of them if the two perimeters are equal), delete it and take the mirror spatial region of the remaining part with respect to the considered axis. The figure thus obtained has the same area, a shortest total perimeter and is symmetric with respect to the considered axis. Since the axes can be chosen, in the euclidean plane, with an arbitrary orientation, the minimum perimeter figure has to be symmetric with respect to any axis cutting it into two equal area parts, whichever is its orientation. The only figure symmetric with respect to an axis cutting it into two equal area parts, independently from its orientation, is the *circle*.

Now, note that in our discrete framework the concept of axis becomes that of an infinite sequence of cells each one adjacent to its predecessor along the same adjacency direction. This means that there exist only four directions with respect to which to consider the symmetry: the horizontal, the vertical and the two 45 degrees diagonal axes. Therefore, the spatial region with maximum aggregation has to be almost symmetric with respect to these four directions. 'Almost' is due to the fact that it is easy to see that for some number of cells there not exist any spatial region that is symmetric with respect to all the four directions, as the simple example in Figure 2 shows.

It is therefore useful to consider those regular polygons definable in our discrete framework, which have the above discussed symmetries. The ones with smaller number of sides are the *square*, the *rhomb* and the *octagon*. An example of these figures drawn on an  $8 \times 8$  grid is given in Figure 3.

Let us denote with  $k$  the number of black cells composing a given region. To build a regular square  $S_s$  of side  $s$  it has to be:

$$k = s^2 \quad (1)$$

while to build a regular rhomb  $R_s$  of side  $s$  it has to be:

$$k = 2s^2 - 2s + 1 \quad (2)$$

and to build a a regular octagon  $O_s$  of side  $s$  it has to be:

$$k = 7s^2 - 10s + 4 \quad (3)$$

with the assumption that corner cells count one for the two sides to which they belong. It is not hard to see that for such regular figures, the aggregation factor holds:

$$\begin{aligned} agg(S_s) &= \frac{8s^2 - 12s + 4}{8} \\ agg(R_s) &= \frac{16s^2 - 32s + 16}{8} \\ agg(O_s) &= \frac{56s^2 - 108s + 52}{8} \end{aligned} \quad (4)$$

We can now explicit  $k$  with respect to  $s$  in formulae (1)–(3) and substitute it in the corresponding formulae (4). We thus obtain the aggregation factor for the three figures with  $k$  black cells:

$$agg_{SQUARE}(k) = \frac{8k - 12\sqrt{k} + 4}{8} \quad (5)$$

$$agg_{RHOMB}(k) = \frac{8k - 8\sqrt{2k-1}}{8} \quad (6)$$

$$agg_{OCTAGON}(k) = \frac{8k - 4\sqrt{7k-3}}{8} \quad (7)$$

which hold for those integer values of  $k$  for which in formulae (1)–(3)  $s$  is integer. Namely, we have that it has to be  $k \in \{x \mid \sqrt{x} \in \mathbb{N}\}$  for (5),  $k \in \{x \mid \sqrt{2x-1} \in \mathbb{N}, x \in \mathbb{N}\}$  for (6), and  $k \in \{x \mid \sqrt{7x-3} \in \mathbb{N}, x \in \mathbb{N}\}$  for (7). If we assume for continuity that for all  $k \in \mathbb{N}$  the above relations hold, then it can be shown that:

$$agg_{OCTAGON}(k) > agg_{RHOMB}(k) > agg_{SQUARE}(k)$$

We can then conclude that the octagon has the property to be the most aggregated figure among the three proposed. The average adjacency number of a cell belonging to an octagon made up of  $k$  black cells is:

$$\overline{adj}_{OCTAGON}(k) = \frac{8k - 4\sqrt{7k-3}}{8k}$$

that tends rapidly to 1 as soon as  $k$  increases. For typical values of percentage of black in a standard image space  $1024 \times 1024$ , that is for values of  $k$  ranging from 200,000 to 500,000 (i.e., existence probability comprised approximately between 0.2 and 0.5), the respective average adjacency of a cell belonging to the octagon ranges in:

$$0.9999975 < agg_{OCTAGON}(k) < 0.999999$$

and this means that from a practical point of view it will be good enough simply to normalize the aggregation factor with respect to the number of black cells belonging to the spatial region to obtain a meaningful measure.

Resuming, we redefine the *coefficient of aggregation*  $\alpha(\mathcal{I}_k)$  of a spatial region  $\mathcal{I}_k$  made up of exactly  $k$  cells as:

$$\alpha(\mathcal{I}_k) = \frac{agg(\mathcal{I}_k)}{k}$$

Note that with this simplified definition of  $\alpha(\mathcal{I}_k)$ , no spatial region having a finite number of cells with  $\alpha(\mathcal{I}_k) = 1$  can be produced. From such a definition, the following immediately descends:

1. If  $\mathcal{I}_k$  is constituted of  $k$  single-points, then of course each one of these cells has an adjacency number equal to zero, and so it will be  $agg(\mathcal{I}_k) = 0$ , from which  $\alpha(\mathcal{I}_k) = 0$ ;
2.  $\alpha(\mathcal{I}_k)$  has a maximum whenever  $\mathcal{I}_k$  coincides with  $\mathcal{I}_k^*$ , that is the most aggregated spatial region for the number of black cells constituting  $\mathcal{I}_k$ ;
3.  $\alpha(\mathcal{I}_k)$  does not depend by the spatial position of the spatial region  $\mathcal{I}_k$ , which is obvious, since of course  $agg(\mathcal{I}_k)$  does not change in consequence of translations.

### 3 Generating spatial regions with the new model

We are now ready to define the new random model. Given an  $n \times n$  space, a random spatial region  $\mathcal{I}$  is generated with respect to two distinct parameters:  $\alpha(\mathcal{I})$ , representing the coefficient of aggregation of the spatial region and  $\rho(\mathcal{I})$ , representing the probability for a cell to be black. An algorithmic way to produce such an expected spatial region is the following: as first step, we give the colour black to each cell belonging to the spatial region (starting for example from the left uppermost cell and proceeding row by row) with probability  $\rho(\mathcal{I})$ . After that, we calculate the coefficient of aggregation of the generated spatial region and we compare it with  $\alpha(\mathcal{I})$ . Then we randomly adjust the spatial region moving its black cells in a way that aims to decrease or to increase the coefficient of aggregation, tending towards  $\alpha(\mathcal{I})$ . This can be done in the following way: we choose at random a black cell  $b$  and a white cell  $w$ , counting the number of their adjacent black cells (without considering  $b$  as an adjacent black cell for  $w$  in the case  $b$  and  $w$  are adjacent). There are three possibilities:

$adj(w) < adj(b)$  : in this case a swap between the cells will reduce the coefficient of aggregation of the spatial region;

$adj(w) = adj(b)$  : in this case a swap between the cells will not change the coefficient of aggregation of the spatial region;

$adj(w) > adj(b)$  : in this case a swap between the cells will increase the coefficient of aggregation of the spatial region.

Note that if we are trying to produce a spatial region with higher coefficient of aggregation and we swapped cells only in the case  $adj(w) > adj(b)$ , we could reach a point where no cell could be further moved but the desired  $\alpha(\mathcal{I})$  is not reached. This happens since each 'island' of cells large enough behaves like an 'attractor' tending to assume a stable configuration having a locally maximal degree of aggregation. Therefore we swap cells even if their adjacency number is the same. We will stop the moulding process as soon as the expected  $\alpha(\mathcal{I})$  is reached by the coefficient of aggregation of the generated spatial region. Note that it may be possible that the requested value of  $\alpha(\mathcal{I})$  is not reachable with the number of cells produced. This may happen in two cases:

1. the requested value of  $\alpha(\mathcal{I})$  is greater than the current one and cannot be reached given the simplified definition of the coefficient of aggregation;
2. the requested value of  $\alpha(\mathcal{I})$  is lower than the current one and cannot be reached due to the limited  $n \times n$  space available to arrange the given number of cells. Note that in a limited  $n \times n$  space the cells on the corners of the image space can reach at most an adjacency number  $adj(p) = 3/8$ , while the cells along the sides can reach at most an adjacency number  $adj(p) = 5/8$ . This is correct since if we gave to a cell in a corner of the space with three adjacent black cells an adjacency number equal to 1, this would make the coefficient of aggregation not invariant with respect to translations, and furthermore the cells along the border of the space would result attractive during the moulding process of the spatial region.

However, these are only theoretical limitations, since in an applicative context interesting values of the coefficient of aggregation are far from to the above two bounds. In the following, a pseudo-Pascal version of the algorithm is given:

```

CREATE_SPATIAL_REGION(RHO,ALPHA)
/* Construct a spatial region having existence probability
   RHO and coefficient of aggregation ALPHA */
cell array REGION[1..N][1..N];
/* REGION contains the array of cells */

float AGG;
cell B, W; /* a cell type is a couple of coordinates */
begin
GENERATE_REGION(RHO,REGION);
/* create a random region where each cell
   has existence probability RHO */
AGG:=CALCULATE_AGGREGATION(REGION);
/* determine the coefficient of aggregation
   of the generated region */
if (ALPHA>AGG) then /* region has to be aggregated */
  while (ALPHA>AGG) begin
    repeat
      B:=RANDOM_BLACK(REGION);
      /* choose at random a black cell */
      W:=RANDOM_WHITE(REGION);
      /* choose at random a white cell */
    until adj(B)≤adj(W);
    EXCHANGE(B,W,REGION);
    /*swap colour of cells B and W */
    AGG:=CALCULATE_AGGREGATION(REGION);
  end
else /* region has to be disaggregated */
  while (ALPHA<AGG) begin
    repeat
      B:=RANDOM_BLACK(REGION);
      W:=RANDOM_WHITE(REGION);
    until adj(B)≥adj(W);
    EXCHANGE(B,W,REGION);
    AGG:=CALCULATE_AGGREGATION(REGION);
  end
return(REGION);
end.

```

Note that proceeding in this way we separate the probability for a cell to be black from the degree of aggregation. For example, a low  $\rho(\mathcal{I})$  matched with a high  $\alpha(\mathcal{I})$  will generate few aggregated regions inside the  $n \times n$  space, while on the contrary a high  $\rho(\mathcal{I})$  matched with a low  $\alpha(\mathcal{I})$  will generate a noisy spatial region.

### 4 Experimental results

In this section we provide a gallery of examples showing how the generator of aggregated images works. We limit our attention to meaningful values for the existence probability, as derived from the experience. For graphical reasons, image space is limited to  $64 \times 64$ . Table 1 contains summarizing parameters for the examples shown in Figures 4-10:

Fig.	$\rho(\mathcal{I})$	$\alpha(\mathcal{I})$	Number of black cells	Initial value of $\alpha(\mathcal{I})$	Final value of $\alpha(\mathcal{I})$	Number of exchanges
4	0.2	0.8	736	0.165761	0.800611	4192
5	0.2	0.9	830	0.205723	0.900000	6553
6	0.2	0.95	860	0.194477	0.950000	19848
7	0.5	0.8	2019	0.480560	0.800396	2754
8	0.5	0.95	2070	0.491425	0.950121	13146
9	0.5	0.97	2056	0.490429	0.970008	61597
10	0.8	0.97	3264	0.777420	0.970052	4043

Table 1: Summarizing parameters for the proposed figures

Finally, Figures 11-13 contain charts showing the boundary values of the coefficient of aggregation for the studied

values of existence probability. Experiments show that the coefficient of aggregation tends to be stable after a number of exchanges in the order of 20,000. Note that each chart shows with a continuous line an aggregation process on the random spatial region and with a dashed line a disaggregation process on the same region.

## 5 Conclusions

In this paper we have proposed a random model for spatial regions that, starting from the traditional pixel independent model, provides the possibility to mould the generated set of points until a desired level of aggregation is reached. We showed how the flexibility of this new model is useful in an applicative context, since it permits to focus on subclasses of spatial regions having similar spatial and quantitative points' distribution.

Future work will be focused mainly to individuate, given a number of points, lower and upper bounds for the associate aggregation factor.

## References

- [1] C.H. Ang and H. Samet, A Fast Quadtree Normalization Algorithm, *Pattern Recognition Letters*, (15), pp. 57-63, January 1994.
- [2] C. Faloutsos, Analytical Results on the Quadtree Decomposition of Arbitrary Rectangles, *Pattern Recognition Letters*, (13) 1, pp. 31-40, January 1992.
- [3] Y. Manolopoulos, E. Nardelli, G. Proietti and M. Vassilakopoulos, On the Creation of Random Quadtrees by Using a Branching Process, *Image Vision and Computing*, under revision.
- [4] E.Nardelli and G.Proietti, Probabilistic Models for Images and Quadtrees: Differences and Equivalences, TR-402, Institute for Systems Analysis and Informatics, C.N.R., Roma, January 1995.
- [5] C. Puech and H. Yahia, Quadtrees, Octrees, Hyperoctrees: a Unified Analytical Approach to Tree Data Structures Used in Graphics, Geometric Modeling and Image Processing, in *Proc. of the Symposium on Computational Geometry*, Baltimore, Maryland, June 1985, pp.272-280.
- [6] H. Samet, *The Design and Analysis of Spatial Data Structures*, Addison-Wesley, Reading, MA, 1990.
- [7] M. Vassilakopoulos and Y. Manolopoulos, Analytic Results on Quadtree Average Storage Requirements, in *Proc. of the 5th International Conference on Computer Analysis of Images and Patterns (CAIP93)*, IAPR, Springer-Verlag, Lecture Notes in Computer Science #719, pp.41-48, Budapest, September 1993.
- [8] M. Vassilakopoulos and Y. Manolopoulos, On Random Models for Analyzing Region Quadtrees, *Pattern Recognition Letters*, accepted.

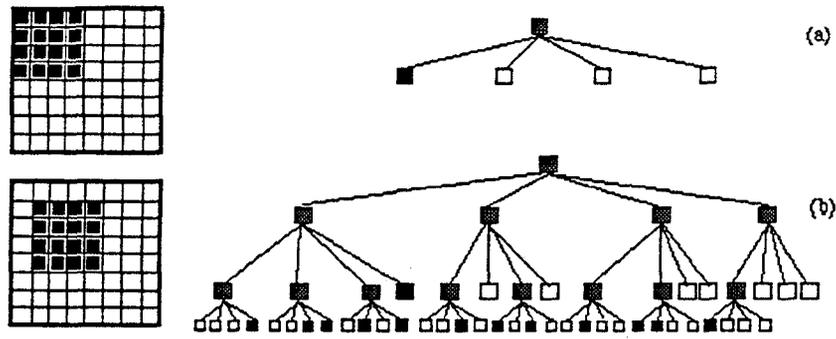


Figure 1: The same region can generate a dense (a) or a bushy (b) quadtree.

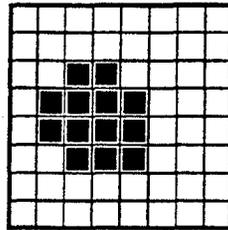


Figure 2: A non-symmetric spatial region.

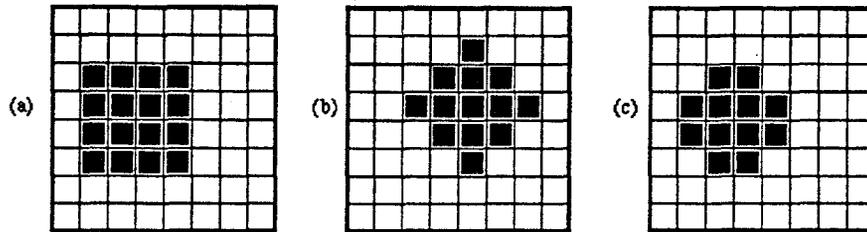


Figure 3: A square of side 4 (a), a rhomb of side 3 (b) and an octagon of side 2 (c) on an  $8 \times 8$  grid.

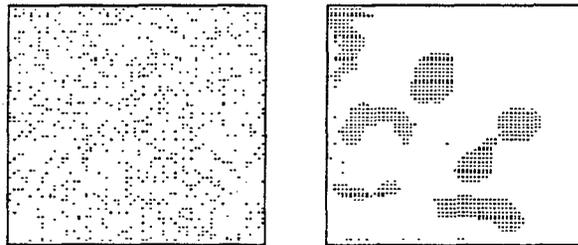


Figure 4

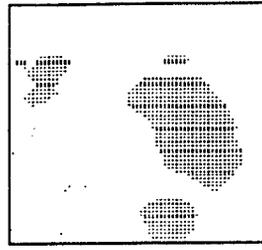
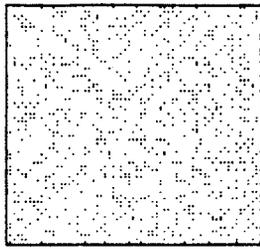


Figure 5

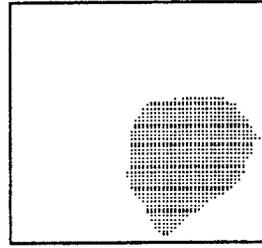
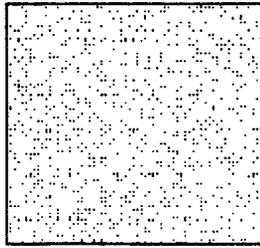


Figure 6

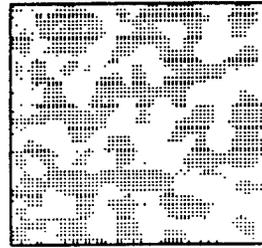
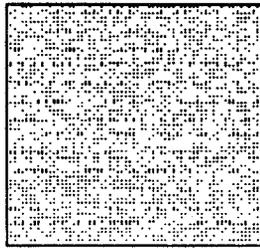


Figure 7

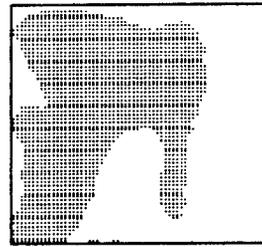
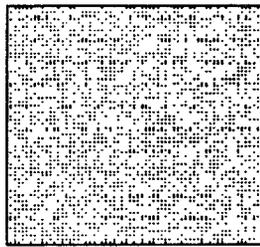


Figure 8

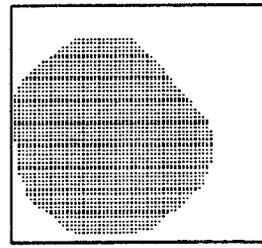
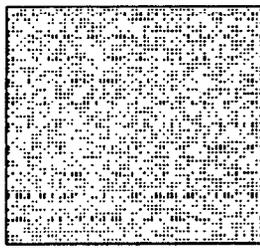


Figure 9

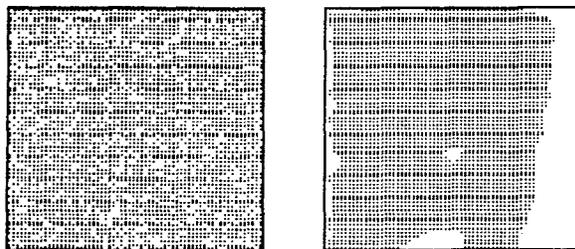


Figure 10

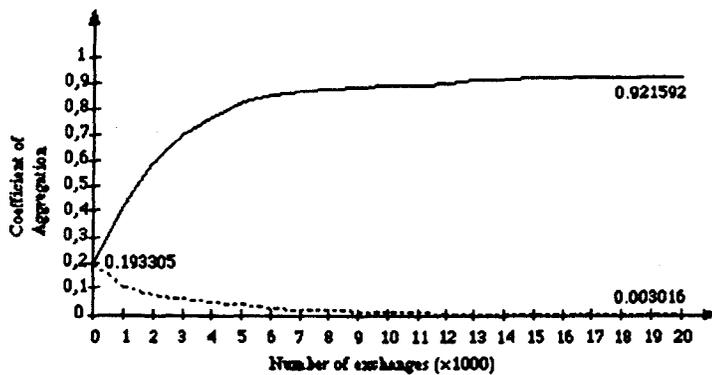


Figure 11: Boundary conditions for  $\rho(I) = 0.2$

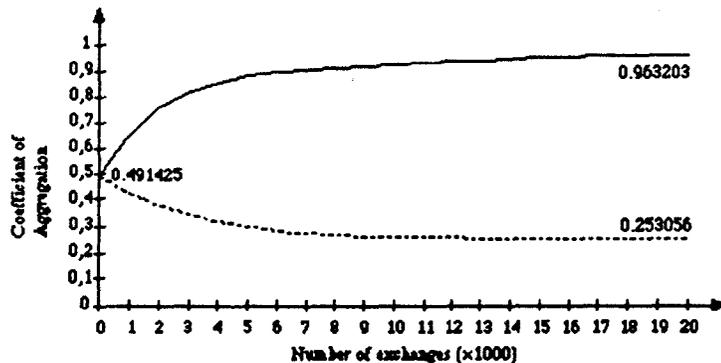


Figure 12: Boundary conditions for  $\rho(I) = 0.5$

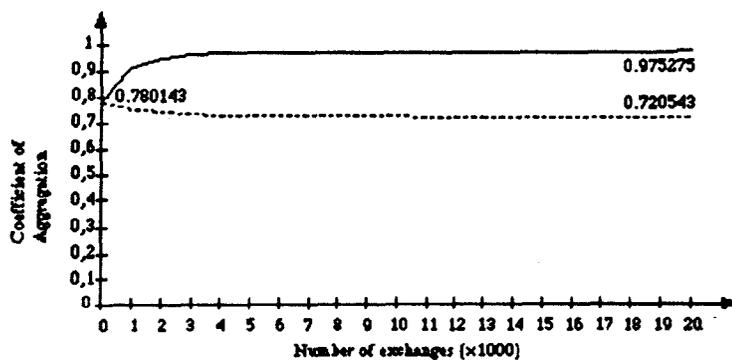


Figure 13: Boundary conditions for  $\rho(I) = 0.8$