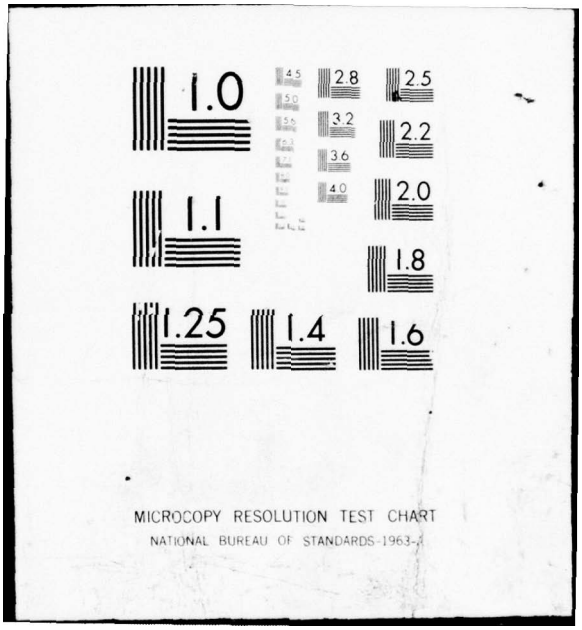
AD-A050 100 MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER F/G 9/2
MOSAIC MODELS FOR IMAGE ANALYSIS AND SYNTHESIS.(U)
NOV 77 N AHUJA AFOSR-77-3271
UNCLASSIFIED TR-607 AFOSR-TR-78-0127 NL

| OF |
AD
A050 100

END
DATE
FILMED
3-78
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

COMPUTER SCIENCE TECHNICAL REPORT SERIES

Approved for public release;
distribution unlimited.

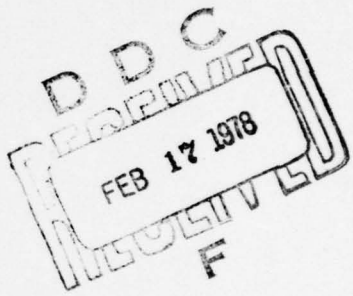
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

TR-607
AFOSR-77-3271

November 1977

MOSAIC MODELS FOR IMAGE ANALYSIS AND SYNTHESIS

Narendra Ahuja
Computer Science Center
University of Maryland
College Park, MD 20742



ABSTRACT

This report is concerned with a class of models for natural images that treat an image as a mosaic, composed of patches generated by a random geometric process. These models are contrasted with conventional statistical image models. Proposed applications of such models in image analysis and synthesis are discussed.

The support of the Directorate of Mathematical and Information Sciences, U. S. Air Force Office of Scientific Research, under Grant AFOSR-77-3271 is gratefully acknowledged, as is the help of Mrs. Shelly Rowe in preparing this paper.

1. Introduction

Real world images very often appear to be piecewise constant. At any level of detail of the scene captured in the image, one can often identify patches that are relatively uniform and have relatively sharp borders. Sometimes these patches are very small compared to the image size, so that, in spite of the sharpness of the inter-patch borders, the image can be regarded as a uniform point process.

The image models currently in use characterize the entire image, or its TV scan, as a stochastic process that best accounts for the gray level variations. Thus, these models are primarily applicable to the subclass of images that are relatively uniform, as discussed above. This report describes a class of "mosaic models" that assume the image to be a spatial arrangement of patches.

In Section 2, we discuss the important properties of the various kinds of statistical image models that have been conventionally used. We subdivide these into two major classes: time series models and random field models.

In Section 3 and Appendices A and B we describe some mosaic models, and contrast their properties with those of conventional models. The implications of these differences for the relative usefulness of the models in dealing with the complexity of natural images are also discussed.

Section 4 discusses the approach we propose to take toward applying mosaic models to image analysis and synthesis, and Section 5 briefly discusses the potential applications of the proposed work.

2. Statistical Image Models

We will divide the types of models used for images into two categories:

2.1 Time Series Models

Time series analysis [1] has been extensively used [2, 3, 12] to study visual textures. The image is TV scanned to provide a one-dimensional series of gray level fluctuations, which is treated as a one-dimensional stochastic process evolving in "time". The future course of the process is presumed to be predictable by knowing enough about its past.

Before summarizing the models, we would like to introduce some often used notation in time series.

Let

$$\cdots\ Z_{t-1}\ Z_t\ Z_{t+1}\ \cdots$$

be a discrete time series where Z_i is the value of the random variable Z at time i. We denote the series by [Z].

Let μ be the mean of [Z], called the "level" of the process.

Let $[\tilde{Z}]$ denote the series of deviations about μ, i.e.,

$$\tilde{Z}_i = Z_i - \mu$$

Let [a] be a series of outputs of a white noise source, with mean zero and variance σ_a^2.

Let B be the "backward" shift operator such that

$B\ \tilde{Z}_t = \tilde{Z}_{t-1}$; hence

$B^m\ \tilde{Z}_t = \tilde{Z}_{t-m}$;

and let ∇ be the "backward" difference operator such that

$$\nabla\tilde{Z}_t = \tilde{Z}_t - \tilde{Z}_{t-1} = (1-B)\tilde{Z}_t;$$

hence $\nabla^m\tilde{Z}_t = (1-B)^m\tilde{Z}_t$

The dependence of the current value of the random variable $\tilde{Z}$, $\tilde{Z}_t$, on the past values of $\tilde{Z}$ and a is expressed in different ways, and this gives rise to several different models [12].

(a) Autoregressive Model:

In this model the current $\tilde{Z}$-value depends on the previous p $\tilde{Z}$-values, and on the current noise term:

$$\tilde{Z}_t = \phi_1\tilde{Z}_{t-1} + \phi_2\tilde{Z}_{t-2} + \ldots + \phi_p\tilde{Z}_{t-p} + a_t \qquad (1)$$

If we let

$$\phi_p(B) = 1-\phi_1B-\phi_2B^2-\ldots-\phi_pB^p$$

then (1) becomes

$$[\phi_p(B)]\ (\tilde{Z}_t) = a_t$$

$[\tilde{Z}]$, as defined above, is known as the autoregressive process of order p, and $\phi_p(B)$ as the autoregressive operator of order p. The name "autoregressive" comes from the model's similarity to regression analysis, and the fact that the variable $\tilde{Z}$ is being regressed on previous values of itself.

(b) Moving Average Model:

In (a) above, $\tilde{Z}_{t-1}$ can be eliminated from the expression for $\tilde{Z}_t$ by substituting

$$\tilde{Z}_{t-1} = \phi_1\tilde{Z}_{t-2} + \phi_2\tilde{Z}_{t-3} + \ldots + \phi_p\tilde{Z}_{t-p-1} + a_{t-1}$$

This process can be repeated to yield eventually an expression for $\tilde{z}_t$ as an infinite series in the a's.

The moving average model allows a finite number q of previous a-values in the expression for $\tilde{z}_t$. This explicitly treats the series as being observations on linearly filtered Gaussian noise.

Letting

$$\theta_q(B) = 1-\theta_1 B-\theta_2 B^2- \ldots -\theta_q B^q,$$

we have

$$\tilde{z}_t = [\theta_q(B)](a_t)$$

as the moving average process of order q.

(c) Mixed Model:

To achieve greater flexibility in fitting of actual time series, this model includes both the autoregressive and the moving average terms. Thus

$$\tilde{z}_t = \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} +\ldots+ \phi_p\tilde{z}_{t-p} + a_t-\theta_1 a_{t-1}-\theta_2 a_{t-2}-\ldots-\theta_q a_{t-q}$$

$$\text{i.e., } [\phi_p(B)](\tilde{z}_t) = [\theta_q(B)](a_t) \qquad (2)$$

In all the three models just mentioned, the process generating the series is assumed to be in equilibrium about a constant mean level. Such models are called stationary models.

There is another class of models called non-stationary models, in which the level μ does not remain constant. The series involved may, nevertheless, exhibit homogeneous behavior when the differences due to level-drift are accounted

for. It has been shown [1] that such a behavior may be represented by a generalized autoregressive operator $\hat{\phi}(B)$, in which one or more of the zeroes of the polynomial $\hat{\phi}(B)$ (i.e., roots of the equation $\hat{\phi}(B) = 0$) is unity.

Thus

$$\hat{\phi}(B) = \phi(B) \cdot (1-B)^d$$

where $\phi(B)$ is a stationary operator.

The general model for a homogeneous nonstationary process, therefore, is of the following form:

$$\hat{\phi}_p(B)\tilde{z}_t = \phi_p(B)(1-B)^d\tilde{z}_t = \theta_q(B)a_t$$

i.e., $$\phi_p(B)w_t = \theta_q(B)a_t$$

where $$w_t = (1-B)^d\tilde{z}_t = \nabla^d\tilde{z}_t \qquad (3)$$

Comparing (3) with (2), we see that the homogeneous nonstationarity can be accounted for by appropriately reweighting the differences.

A time series may show a repetitive pattern of periods of similar characteristics. For example, in the TV scan of an image the intervals corresponding to rows will have similar characteristics. The following generalized model [12] incorporates the presence of such "seasonal effects" in the time series:

$$\phi_p(B)\ \phi'_p(B^s)\ \nabla^d\ \nabla_s^D\ \tilde{z}_t = \theta_q(B)\ \theta'_q(B^s)\ a_t$$

where s is the seasonal period,

$$\nabla_s = (1-B^s),$$

and $\phi'_p(B^s)$ and $\theta'_q(B^s)$ are polynomials in B^s, of order p and q, respectively.

By studying the statistical properties of a given texture, e.g., its autocorrelation function, etc., McCormick and Jayaramamurthy [12] have made a choice of the best fitting model from among those described above. They also use the same information to estimate the required set of parameter values, and generate synthetic texture using the model.

In an earlier paper Whittle [21] pointed out the difficulty of using time series for spatial processes. The problem is that in a two-dimensional process, the dependence of a point extends in all directions, and there is no direct way to map the two-dimensional grid points onto a series such that the original dependence is preserved, although it is unilateral (depends only on the past values). Whittle analyzes the problem in two steps:

(1) Consider a series in which a point depends upon both the preceding and the following points. For example, $\tilde{Z}_t$ may depend on $\tilde{Z}_{t-1}$ and $\tilde{Z}_{t+1}$. Whittle obtains a set of constraints on $[\tilde{Z}]$, and the corresponding bilateral autoregression operator, such that the influence of the future values vanishes.

(2) Now consider a two-dimensional grid. The simplification amounts, in terms of Fig. 1, to expressing the

value at the O as a function of the values at the upper-case X's, when, in fact, it depends upon all the X's, both upper and lower-case. Whittle argues that this is a much more complicated process than the one-dimensional case in (1), and hence, there is no escaping the explicit introduction of dependence in all directions.

He derives the conditions to be satisfied by the parameters optimizing the reduced models resulting in (1) and (2). While it is already difficult to obtain the appropriate values in case (1), case (2) is still worse, in addition to being less accurate. He points out that the problem is due to the fact that in two dimensions, unlike one dimension, the unilateral representation of a finite autoregression is not a finite autoregression. The real usefulness of the unilateral representation is that it suggests a simplifying change of parameters. For most two-dimensional models, however, the appropriate transformation, even if evident, is so complicated that nothing is gained by performing it.

One would like to try to capture as much of the two-dimensional dependence as possible without getting into the analytical problems due to bilateral dependence. Tou et al. [2] have done this by making a point depend on its upper and left neighbors, as shown in Figure 2; the O depends only upon the X's. For such a case, the autoregressive process of order (q,p) is

$$\tilde{z}_{ij} = \phi_{01}\,\tilde{z}_{i,j-1} + \phi_{10}\,\tilde{z}_{i-1,j} + \phi_{11}\,\tilde{z}_{i-1,j-1} + \ldots + \phi_{qp}\,\tilde{z}_{i-q,j-p};$$

the moving average process of order (q,p) is

$$\tilde{z}_{ij} = a_{ij} - \theta_{01} a_{i,j-1} - \theta_{10}\,a_{i-1,j} - \theta_{11}\,a_{i-1,j-1} - \ldots - \theta_{qp}\,a_{i-q,j-p};$$

and the two-dimensional mixed autoregressive/moving average process is

$$\tilde{z}_{ij} = \phi_{01}\,\tilde{z}_{i,j-1} + \phi_{10}\,\tilde{z}_{i-1,j} + \phi_{11}\,\tilde{z}_{i-1,j-1} + \ldots + \phi_{qp}\,\tilde{z}_{i-q,j-p}$$

$$+ a_{ij} - \theta_{01}\,a_{i,j-1} - \theta_{10}\,a_{i-1,j} - \theta_{11}\,a_{i-1,j-1} - \ldots$$

$$- \theta_{rs}\,a_{i-r,j-s}$$

If the coefficients of the process satisfy the condition

$$\phi_{mn} = \phi_{m0}\,\phi_{0n}$$

then the process becomes a multiplicative process in which the influence of rows and columns on the autocorrelation is separable. Thus

$$\rho_{ij} = \rho_{i0}\,\rho_{0j}$$

Tou et al. consider fitting a model to a given texture. The choice among the autoregressive, moving average and mixed models, as well as the choice of the order of the process, is made by comparing the behavior of some observed statistical properties, e.g., the autocorrelation function, with that pre-

dicted by each of the different models. For each of the possibly many choices of models, the values of the parameters are determined so as to minimize, say, the least square error in fit.

A comparison of the predictions of autocorrelation function, results of transformations of the series, etc., based upon the model obtained above, with similar properties of the available data can be used to establish its appropriateness, or to suggest desirable modifications in the model, e.g., changing the order, etc.

In a subsequent paper, Tou and Chang [3] use the maximum likelihood principle to optimize the values of the parameters, in order to obtain a refinement of the preliminary model as suggested by the autocorrelation function.

2.2 Random Field Models

The second class of models treats the image as a two-dimensional random field (for a definition of random field, see Rosenfeld and Kak [20]).

Oceanographers [9, 10, 11, 13] have long been interested in the patterns formed by waves on the ocean surface. Longuet-Higgins [9, 10, 11] treats the ocean surface as a random field satisfying the following assumptions:

(a) the wave spectrum contains a single narrow band of frequencies, and

(b) the wave energy is being received from a large number of different sources whose phases are random.

Considering such a random field, he obtains [11] the statistical distribution of wave heights, and derives relations between the root mean square wave height, the mean height of the highest p% of the waves, and the most likely height of the largest wave in a given interval of time.

In subsequent papers [9, 10], Longuet-Higgins obtains an additional set of statistical relations among the parameters describing (a) a random moving Gaussian surface [9], and (b) a Gaussian isotropic surface [10].

Some of the results that he derives are:

(1) the probability distribution of the surface elevation, and that of the magnitude and orientation of the gradient,

(2) the average number of zero crossings per unit dis-

tance along a line in an arbitrary direction,

(3) the average length of contour per unit area,

(4) the average density of maxima and minima per unit area, and

(5) for a narrow spectrum, the probability distribution of the heights of maxima and minima.

All the results are expressed in terms of the two-dimensional energy spectrum up to a finite order only. The converse of the problem is also studied and solved, i.e., given certain statistical properties of the surface, to find a convergent sequence of approximations to the energy spectrum.

The analogy between this work and image processing, and the significance of the results obtained therein, is obvious. Fortunately the assumptions made are also acceptable for images.

Panda [15] uses an analogous approach to analyze background regions selected from Forward Looking InfraRed (FLIR) imagery. He derives expressions for (a) density of border points and (b) average number of connected components in a row of the thresholded picture. There is good agreement between the observed and the predicted values in most cases, for most of the pictures considered. Panda [14] also uses the same model to predict the properties of the pictures obtained by running several edge operators (based on differences of average gray levels) on some synthetic pictures with normally distributed gray levels, and having different correlation co-

efficients. The images are assumed to be continuous-valued stationary Gaussian random fields with continuous parameters.

Hunt [16, 17] points out that stationary, Gaussian modelling of images is an oversimplification. Consider the vector F of the picture points obtained by concatenating them as in a TV scan. Let R_F be the covariance matrix of the gray levels in F. Then according to the Gaussian assumption, the probability density function of F is

$$p(F) = K \exp[-\frac{1}{2}(F-\bar{F})\ R_F^{-1}(F-\bar{F})]$$

where $\bar{F}$ = mean vector of identical, constant gray level components

R_F = covariance matrix

and K = normalizing constant

The stationarity assumption makes $\bar{F}$ a vector of identical components. This means that each point in the image has the same ensemble statistics. Images, however, seldom have a bell-shaped histogram.

A Gaussian model for any set of multivariate data, however, is the only model that is mathematically tractable to any reasonable extent. Hunt [16] proposes a nonstationary Gaussian model which differs from the stationary model only in that the mean vector $\bar{F}$ has unequal components. He shows the appropriateness of this model by subtracting, from each point on the image, its local ensemble average, and showing that the resulting picture fits a stationary Gaussian model

reasonably well, the mean vector $\bar{F}$ being the vector of local ensemble averages.

In work on image restoration, images have often been modelled by a two-dimensional random field with a given autocorrelation function. An autocorrelation function that has been found to be reasonably good [4-8] for a variety of pictorial data is

$$R(\tau_1,\tau_2) = \sigma^2 \exp[-\alpha_1|\tau_1| - \alpha_2|\tau_2|],$$

which is stationary and separable.

In a recent paper, Nahi and Jahanshahi [18] model the image as a background statistical process combined with a set of foreground statistical processes, each replacing the background in the regions occupied by the objects of the category which it is assumed to characterize. In estimating the boundaries of horizontally convex objects on a background in noisy binary pictures, Nahi and Jahanshahi assume that the two kinds of regions in the picture are formed by two statistically independent stationary random processes with known (estimated) first two moments. The end-points of the intercepts of the given object on successive rows are assumed to form a first order Markov process.

Thus, using the notation

$b(m,n)$ = gray level at the nth column of the mth row

$\gamma(m,n)$ = a binary function carrying the boundary information

b_b = a sample gray level from the background process,

b_o = a sample gray level from the object process, and

ν = a sample gray level from the noise process,

the model allows us to write

$$b(m,n) = \gamma(m,n)\, b_o(m,n) + [1-\gamma(m,n)]\, b_b(m,n) + \nu(m,n)$$

where γ incorporates the Markov constraints on the object boundaries.

In a theoretical treatment, Wong [22] discusses the characterization of second order random fields (having finite first and second moments) from the point of view of their possible use in representing images. He considers various properties of a two-dimensional random field, and their implications in terms of its second order properties. Some of the results he obtains are as follows:

(1) There is no continuous Gaussian random field of two dimensions (or higher dimensions) which is both homogeneous and Markov (degree 1).

(2) If the covariance function is invariant under translation as well as rotation, then it can only depend upon the Euclidian distance. The second order properties of such fields (Wong calls them homogeneous) are characterizable in terms of a single one-dimensional spectral distribution.

Wong generalizes his notion of homogenity to include random fields that are not homogeneous, but can be easily transformed into homogeneous fields. Even this generalized class of fields is no more complicated than a one-dimensional stationary process.

3. Mosaic Image Models

We briefly reviewed, in Section 2, various types of models that have been used in earlier work on image modelling. In this section we consider another class of models which, we think, provides an alternative that may prove to be useful for a large class of natural images.

Schachter and Ahuja [19] (reproduced here as Appendix A) discuss several processes that could give rise to visual spatial patterns. Many of these processes, although very likely to be present in nature, and hence to give rise to natural images, are too complex to be mathematically tractable. [19] discusses, at length, the description, specification, generation and properties of the models based on some of the relatively less complex processes, and brings out the differences between the class of models being proposed and those reviewed in Section 2. It also provides examples of the patterns generated by some of the processes. An additional type of process, based on a random walk, is described in Appendix B.

Various similarities and differences among the mosaic, time series, and random field models will now be discussed.

(1) Mosaic models describe images by specifying geometrical processes that may have generated the visual pattern under consideration. Such a constructive description, therefore, inherently encompasses the specification of all the information about the pattern. One may extract from the model

as much information as desired, e.g., autocorrelation properties which may not be unique to the image. For example, characterization of a pattern in terms of its autocorrelation properties ignores any phase information.

(2) Time series models allow the current value of an image point to depend on a finite number of previous values. There is, thus, an inherent assumption in the definition of the model about the Markovianity of the data. While it is a different issue how useful the model could still be in practice, such an assumption places a definite theoretical restriction on the generality of the model. Both the random field and mosaic models are free of such a restriction.

(3) Images are inherently two-dimensional and hence should be treated as *such*.

The time series model clearly fails to meet this requirement. It allows a point to depend on, at best, only a part of its neighborhood. A time series model also cannot make use of the rich class of two-dimensional features, e.g., shape and orientation of subpatterns, edge density, connectedness of components, etc., which seem to play an important role in human perception of images. Some of these features have one-dimensional counterparts which could, in principle, be used. But they are much less useful because of their lesser semantic relevance. The random field and mosaic models, on the other hand, are two-dimensional models. Longuet-Higgins [9, 10, 11] and Panda [14, 15] provide examples of the analysis of two-dimensional features of random fields. Some

two-dimensional features of mosaic models are presented by Schachter and Ahuja [19].

(4) In time series modelling the choice of the model is based upon a qualitative assessment of the autocorrelation function. The order of the underlying process is guessed, to begin with, and then iteratively improved until a set of parameter values is found that, along with the chosen order of the model, predicts an autocorrelation function sufficiently close to the observed one [3]. Thus the process of model specification involves some amount of trial and error.

The problem of fitting random field models to real images does not seem to have been given much thought, although different models have been proposed as described in Section 3. Therefore, we do not have examples to show the amount of difficulty involved in the process. Considering the similarity in the nature of descriptions of the data according to the two models, however, the procedures for fully specifying a random field model for a given image may be expected to be similar in approach, and hence in difficulty, to those used by the time series models.

Furthermore, the model arrived at in order to obtain a good fit of properties such as autocorrelation may, in fact, turn out to be worse than expected, due to the fact that the characterization of the image in terms of correlation properties ignores phase information, as pointed out in (1) above.

It may be observed that in both the time series and ran-

dom field models the complexity of modelling is very unevenly shared by the two levels of (a) model selection, and (b) parameter evaluation. There is only a limited choice about the type of model to be selected, and the specification of the chosen model is the major part of the modelling process. The variety of natural images must, therefore, be represented by the assignment of values to the parameter set of the model.

Mosaic models, on the other hand, are much richer in variety, and each of these models is simpler to specify, as compared to the time series and random field models. The complete process of modeling thus gets more evenly split into the two steps. This should have the effect of reducing the size of the search space when fitting models to a given image.

(5) Natural visual patterns can often be characterized by a repetitive arrangement of certain subpatterns, according to a set of rules. The subpatterns, recursively, may be patterns of smaller extent, but of independent complexity (busyness, entropy, fineness). A small set of relatively less busy subpatterns with sharp borders may give the image a patchy appearance, whereas a large set of busy patterns may give rise to a finer, textured pattern.

One reason why mosaic models may be more appropriate for complex natural images is that they provide a hierarchical character to the problem of image modelling. We will illustrate this through a one-dimensional example.

Consider the function of one variable shown in Figure 3. Clearly, any global, closed form characterization of such a

functional form is going to be complicated because of the lack of well-behavedness of the function at the points of non-differentiability. However, if we realize that the function admits a set of simple, local descriptions, we can characterize it in two steps, as follows:

(a) describe the analytically simpler components (straight lines) individually, and

(b) specify the arrangement of these components that defines the entire functional form.

The hierarchy of such a piecewise decomposition may, in general, consist of any number of levels. The patterns may be recursively decomposed into components until they are easy to describe. For example, the one-dimensional pattern of Figure 4a can be described as formed of subpatterns of the type shown in Figure 4b, which, in turn, can be described as formed of the subpatterns of the type shown in Figure 4c, which are trivial to specify. Clearly, an image can also have more than one kind of component without any added complexity in the hierarchical description. Figures 3 and 4a provide one-dimensional counterparts of the types of images that have been called patchy above.

Mosaic models extend the hierarchical approach to two dimensions. The arrangement of the components, possibly of more than one type, in the mosaic is often specified statistically.

For the nonpatchy class of patterns, both the time

series and random field models may not turn out to be very complex. Although one could also attempt to apply these models to images with regularly shaped patches having relatively sharp borders, such an approach is likely to defy an easy analysis, and is likely to provide a complex model. For example, the order of the resultant time series model may be very high in order to incorporate enough information about the magnitudes of gray level jumps across the patch borders. Clearly, although the interior patch-points do not contain much information, they do increase the order of the model, making it more complex and computationally more expensive.

Not surprisingly, in view of the above observations, the time series model has been used [2, 3, 12] only for those images that are relatively well suited for such an approach, as pointed out above. The work on random field modelling, for the most part, has been confined to suggesting different theoretical models, and has payed little attention to actually fitting the models to real images. The recent work of Panda [14, 15] also uses images that fall under the nonpatchy category.

It may be seen that an image from the patchy category can be transformed into a picture of the nonpatchy category by sampling it sufficiently coarsely. Since this transformation should not change the structure of the image, the model should still be valid with a different set of parameter values. The validity of the choice of a mosaic model thus

appears to be insensitive to scale changes, and mimics the underlying generating process of the image so as to incorporate as much of the detail as is captured in the image to be modelled.

Under certain conditions the mosaic models and the random field models may produce similar patterns. For example, a random field model may fit certain coarsely sampled (dense) mosaics.

(6) Mosaic models are likely to be intuitively more meaningful. A pattern corresponding to a specified model, and the implications for it of the variations in parameter values, may be easier to visualize in case of the mosaic models than the others.

4. Proposed Research

Considering the potential advantages of mosaic models over conventional models as discussed in Section 3, we propose to explore the feasibility of using mosaic models for natural images, and to develop a set of computer programs (hereafter called IMP, for Image Modelling Package) that accepts a homogeneous, isotropic visual pattern as an input, and provides as output a concise description of the image in terms of a completely specified model out of those known to IMP. If none of the known models provides a sufficiently good fit to the input image, a failure may be reported.

At present, we plan to confine ourselves to homogenous, isotropic images, but extensions to other classes of images will be considered for later investigation.

We intend to divide the research into three phases, as follows:

4.1 Choice of Models

Currently, we have in mind the following two broad classes of mosaic models:

(a) Cell Structure Models:

Cell structure models tessellate the plane into nonoverlapping cells (for details, see [19]). Some examples of cell structure models, namely, the Poisson line, occupancy, rotated checkerboard, and rotated hexagon models, are discussed in [19]. In addition, we propose here a random walk model, which is described in Appendix B.

(b) Bombing Models:

Bombing models randomly drop certain geometrical figures on the plane. The area covered by the bombs is colored differently (for further details see [19]). Different kinds of geometrical figures give rise to different kinds of bombing models. We may include the following types of bombs for consideration: circles, ellipses, line segments, rectangles, and squares. We may also consider bombs of a given shape whose sizes are governed by some probability distribution. For example, the radii of the circular bombs, sides of the square bombs, lengths of the linear bombs, etc., can be chosen from a specified distribution over a certain range.

The geometrical processes mentioned in both (a) and (b) are primarily based on their intuitive appeal and expected potential to give rise to interesting patterns. We may find other processes worth considering during the course of the research. Their ultimate selection for use in modelling images, however, is critically dependent on their analytical tractability, as discussed below.

4.2 Analysis of Models

In the second phase of the research we intend to carry out a mathematical analysis of some of the properties of the models chosen for consideration in (1) above. We seek answers to the following kind of questions about the patterns generated by the models:

(a) What is the probability that a pair of points dis-

tance d apart will fall in regions of colors i and j?

(b) What is the edge density per unit area?

(c) What is the expected number of connected components of color i?

(d) What is the expected length of an intercept of the region of color i on a randomly positioned and oriented straight line?

(e) What is the nature of the autocorrelation function?

(f) How do various properties of the pattern change when it is sampled at varying degree of coarseness?

(g) Consider a chain of sample points extracted from the pattern according to a certain scheme, say a chain of equidistant points. Can we identify the one-dimensional stochastic process defined by the chain? Can we describe some of its properties, e.g., autocorrelation function, single-step transition probabilities, expected length of runs of color i, etc.?

(h) What is the expected area of a connected component of color i? Or, what is the total area having color i? Note that in case of cell structure models this information is implicit in the answer to question (c) above, since from the model we already know the stationary probability vector of the colors. This is because the cells are nonoverlapping. However, in the case of bombing models this question

involves a different property.

Some of the features referred to by these questions may be difficult to answer for some of the models because of their mathematical intractability. We would like to obtain as large a set of features for each model as we can. We would also like to explore the possibility of finding the probability distributions of some of the features. Some of the properties that are not mathematically tractable can be empirically estimated.

Based upon the geometrical process underlying a given model, we may be able to visualize some properties of the model that may distinguish the pattern generated by it from those generated by the other models. We may then look for a *dimensionless* measure of such a property and its probability distribution. An example of such a property might be the total border-length per unit area, and the corresponding dimensionless measure might be the square of the border-length per unit area. We will call such a measure an "applicability measure" of the model.

4.3 Application to Image Modelling

Some of the features described in (2), e.g., those dealing with connected components, border per unit area, transition probabilities, etc., concern the interaction among various regions, and not the information within them. These features may be easier to compute on transformed images where each of the regions is given a single gray level.

Our major tool in discovering the appropriate model for the image is the observation of a variety of features on the image. We have two approaches in mind from the point of view of subdividing the modelling process into (a) selecting the type of model, and (b) evaluating its parameters, as described below:

(a) The first approach separates the choice of the model from its parameter specification. An applicability measure, which is a dimensionless quantity, is computed on the image for each of the models. Knowing the distributions of these measures for the various models, we determine the level of confidence with which each model represents the input image. The model having the maximum confidence may be assumed to provide the best description of the image. If none of the models is applicable with the desired degree of confidence, a failure may be reported.

The applicability measure used reflects an abstraction of the image that the model presents. Therefore, its choice has to be made with great care. We may want to include several important geometrical properties of the image in the definition of the applicability measure, so that the chosen model resembles the image to be modelled in as many ways as possible. Since it is not clear what constitute "important" properties, the choice of the applicability measure may need some experimentation.

(b) In the second approach, we consider a sufficiently large set of predicted features that involve all the para-

meters of a model, and use the corresponding observed values to solve for the parameters. Thus, the problems of the choice of the type of model, and the values of its parameters, are treated simultaneously.

Using the computed values of the parameters, the model is used to predict those of the remaining features which have been found predictable for the model under consideration. The model whose predictions are closest to the observations is assumed to be the most appropriate for the given image. Note that there is no notion of an absolute degree of fit, and the approach always yields an answer. The validity of the model depends upon how important the features used for modelling the images are in characterizing visual patterns.

It may be noted that the properties of a one-dimensional series extracted out of the image, which are fundamental to the time series model, are being used here only as some of the many features. Similarly, the autocorrelation function, which often completely characterizes the pattern in the random field models, is used only as one of the features.

Because of the possibility of identically colored cells being contiguous in cell structured models, or of overlap of the bombed areas in bombing models, it may be very difficult to extract from the image features directly relating to a single, isolated cell or a bomb. We shall therefore avoid the use of any such features whenever possible.

Among the various properties of the models discussed above, the transition probabilities as a function of distance

seem to be very powerful in predicting many properties of the image. We shall now give two examples of using the transition probabilities to derive some of the characteristics of the model. These examples also illustrate the kind of analysis of the models that we want to carry out.

(1) Autocorrelation function:

We will derive the expression for the autocorrelation function of a mosaic in terms of the parameters of the underlying model, and of the transition probabilities. Because of the assumptions of homogenity and isotropicity, the autocorrelation function depends only upon the distance. We will also show that one of the mosaic models, the Poisson line model [19], has an autocorrelation function which, as pointed out earlier, happens to be precisely the function that has been widely and successfully used to model a large class of natural images.

We shall use the following notation:

p_i = Probability that a randomly selected point belongs to region i.

$P_{ij}(d)$ = Probability that a pair of points distance d apart, having one point in a type-j cell, has the other point in a type-i cell.

g_i = The mean gray level of the population of gray levels in the type-i cells.

We do not put any constraints on the nature of the distributions of gray levels in the various types of cells, as this

is not crucial for the following analysis.

From the homogeneity and isotropicity properties, we know that the autocovariance is only a function of the distance d. Thus, by definition of autocovariance function, we have

$$AC(d) = \sum_i p_i g_i \sum_j P_{ji}(d)\, g_j - g^2$$

where

$$g = \sum_i p_i g_i$$

is the expected gray level of a randomly selected point. The autocorrelation coefficient, then, is

$$\rho(d) = \frac{AC(d)}{AC(0)} = \frac{AC(d)}{\sigma^2}$$

where $\sigma^2 = AC(0) = Var(g_i)$

Now, for the Poisson line model, we know that [19]

$$P_{ij}(d) = \begin{cases} p_j + (1-p_j)e^{-kd} & \text{if } j=i \\ p_j(1-e^{-kd}) & \text{if } j\neq i \end{cases}$$

where k is a known constant related to the intensity of the Poisson process. Therefore

$$\rho(d) = \frac{1}{\sigma^2}[\sum_i p_i g_i \{\sum_j P_{ji}(d) g_j\} - g^2]$$

$$= \frac{1}{\sigma^2}[\sum_i p_i g_i \{\sum_{j\neq i} P_{ji}(d) g_j + P_{ii}(d) g_i\} - g^2]$$

$$= \frac{1}{\sigma^2}[\sum_i p_i g_i \{ \sum_{j \neq i} p_j (1-e^{-kd}) g_j + (p_i + (1-p_i)e^{-kd}) g_i \} - g^2]$$

$$= \frac{1}{\sigma^2}[\sum_i p_i g_i \{ \sum_{j \neq i} p_j g_j - e^{-kd} \sum_{j \neq i} p_j g_j + p_i g_i + e^{-kd} g_i - e^{-kd} p_i g_i \} - g^2]$$

$$= \frac{1}{\sigma^2}[\sum_i p_i g_i \{ g(1-e^{-kd}) + g_i e^{-kd} \} - g^2]$$

$$= \frac{1}{\sigma^2}[g(1-e^{-kd}) \sum_i p_i g_i + e^{-kd} \sum_i p_i g_i^2 - g^2]$$

$$= \frac{1}{\sigma^2}[g^2(1-e^{-kd}) \div e^{-kd} \cdot E(g_i^2) - g^2]$$

$$= \frac{1}{\sigma^2}[E(g_i^2) - g^2] e^{-kd}$$

$$= \frac{Var(g_i)}{\sigma^2} e^{-kd}$$

$$= \frac{\sigma^2}{\sigma^2} e^{-kd}$$

$$= e^{-kd}$$

which is the exponential model of autocorrelation.

(2) Expected edge density:

If we let E_{ij} denote the expected absolute difference between the gray levels of two points from regions i and j, we can express the edge value between a pair of points at a distance d, E(d), as follows:

$$E(d) = \sum_i p_i \sum_j E_{ij} P_{ij}(d)$$

$$= \sum_i p_i [P_{ii}(d) E_{ii} + \sum_{j \neq i} P_{ji}(d) E_{ij}]$$

Let G_i be a random variable denoting the gray level of a point within region i; then

$$E_{ii} = E|G_{i_1} - G_{i_2}| \text{ , and}$$

$$E_{ij} = E|G_i - G_j|$$

where G_i, G_{i_1} and G_{i_2} are independent and identically distributed random variables.

The expected total edge value between pairs of points at distance d apart is

$$TE(d) = N\ E(d)$$

The expression for E_{ii} and E_{ij} can be derived for various coloring schemes, using different gray level distributions to color different regions in the image.

The fact that the Poisson line model has an exponential autocorrelation function further suggests that the mosaic models may be realistic in the sense that their properties such as autocorrelation function, which have been conventionally treated as characteristic of a pattern, may be similar to those of natural images. The Poisson line model is one of the more random ones [19] and we may expect other mosaic models not to have exponential autocorrelation functions.

Thus, the mosaic models are expected to give rise to a more general class of patterns, including patterns that do not fit the conventional exponential model of autocorrelation, and hence, cannot be represented by models based on that assumption.

5. Summary and Discussion

We have discussed the conventionally used time series and random field models of images, as well as the proposed mosaic models, and have compared their important structural properties. We have also described our proposed approach to the problem of using the class of mosaic models to represent natural images.

Our references to the uses of mosaic models have alluded to their applications in both synthesizing and analyzing images. A knowledge of the physical process that gave rise to a certain pattern may well determine the geometrical process governing the image structure. For example, knowing that an image shows farms, wild vegetation, a heap of hay, leaves or pebbles on ground, drying mud, etc. provides a strong clue to the kind of mosaic model that will be suitable for the image. On the other hand, a study of the patterns generated by various mosaic models can provide an insight into the relationship between the properties of the models and the patterns they generate. This could also help us to understand better the features of the images that most influence their perception by humans.

We believe that IMP will be able to deal with a variety of complex images, and, more important, that the modular nature of its knowledge about the world will allow for continuous, cumulative growth as more processes and models become understood.

APPENDIX A. MOSAIC MODELS

(Technical Report 549)

APPENDIX B. A RANDOM WALK MODEL

In addition to the cell structure models described in Appendix A, we may also consider the following model, based upon a random walk on a two-dimensional grid. We will call it the "random walk" model. It generates a set of curves, beginning at a randomly chosen set of points.

A Poisson process drops points onto a plane. Each point immediately advances in two of the four principal directions with equal probability. The endpoints of the resulting curve segments then independently begin a random walk on the grid.

The number of endpoints before the random walk begins is related to the intensity of the Poisson process and will be at most twice the number of points originally dropped. Each of the endpoints represents one of the ends of a growing curve, and is classified into one of four possible categories according to the direction of its first step, as mentioned in the previous paragraph. This direction characterizes its random walk and will be called its "characteristic direction".

At any step, the probabilities of a point moving in its characteristic direction, or in any of the remaining three directions, depend only on the last step. A walk ends when it hits any of the borders of the plane. When all of the points have finished their walks, their traces provide a random tessellation of the plane.

Since the direction of the current step of any point depends only on the direction of its previous step, its random

walk describes a first-order Markov process.

A nonzero probability of a point taking a step in the direction opposite to the characteristic direction would lead to a nonzero probability of its

(1) retracing some portion of its own walk, and thus having protruding curve segments with dead ends (Figure 5a), and

(2) intersecting its own walk, and thus having isolated loops or loop-stem pairs (Figure 5b).

Since our ultimate interest is to obtain a tessellation of the plane, both (1) and (2) are undesirable. We will, therefore, have zero probability associated with the event that a point takes a step in the direction opposite to the characteristic direction. The probabilities of moving in any of the remaining three directions will depend only upon the previous step such that a point is never allowed to step back to the position it just visited.

The above description of the random walk model is only a specific illustration of using the random walk process to generate a random tessellation of the plane. A generalized model may allow several options at various steps of the tessellation process. Some of these are discussed below:

(1) We have used random curves advancing in two directions, originating at the given set of starting points located throughout the plane. Instead, the starting points could be distributed along the borders of the plane. The curves re-

sulting from the random walks of these points grow only at one of their ends, the other end being fixed at the border.

It is clear that such a process would provide a much less traversed interior of the plane and, hence, the resulting tessellation will have larger cells away from the border. On an average, approximately one-fourth of the total number of starting points would take their very first step out of the plane. Alternatively, a starting point may be assigned a characteristic direction from the set of only those directions which are not directed out of the plane.

(2) We have allowed only the four principal directions for the steps in the random walk of a point. We can also consider steps in the four diagonal directions. However, in order to avoid undesirable patterns as in (1) and (2) above, we will have to prohibit a point from not only taking a step in the direction opposite to the characteristic direction, but also in the two adjacent directions.

When a given point, on its random walk, meets another point, it can do one of two things:

(a) It continues its walk uninfluenced by the other, or,

(b) Its walk merges with that of the other, to result in a single walk which is characterized by one of the two constituent characteristic directions with equal probability.

The first choice mentioned above is likely to be analytically simpler because of lack of interaction among

different random walks.

Among the models discussed in [19], the random walk model comes closest, in terms of the tessellating process, to the Poisson line model. Following are some of the points of comparison between the two models:

(1) Both are cell structure models.

(2) The random walk model uses randomly located zig-zag curves to tessellate the plane, whereas the Poisson line model uses randomly located and oriented straight lines.

(3) The random walk model uses a Poisson process to determine the locations of the points that grow into curves, whereas the Poisson line model uses a Poisson process in the $r\theta$ plane to choose pairs of $r\theta$ values, each of which describes a line in the tessellation.

(4) The cells obtained by the random walk model need not be convex, unlike the situation in the Poisson line model.

(5) The expected number of cells meeting at a vertex in a Poisson line tessellation is 3 [19]. However, in the random walk model this number depends on the nature of the walks; for example, choices (a) and (b) for the interaction between two walks, as mentioned above, yield different results.

The random walk model embodies a process that appears to be similar to those responsible for propagation of cracks in

mud, of borders of vegetation, of boundaries of regions in a map, etc., and may be a reasonable model for a certain class of natural images. It may, therefore, be useful to be able to characterize the random walk tessellation by the intrinsic properties of its cells and their interaction, and to contrast these with similar features of the other cell structure models, especially the Poisson line model; and to make a similar comparison between random walk colored patterns, natural images, and the patterns generated by the Poisson Line model.

REFERENCES

1. Box, J.E.P. and Jenkins, G. M., Time Series Analysis, Holden-Day, San Francisco, 1970.
2. Tou, J. T. and Chang, Y. S., "An approach to texture pattern analysis and recognition", Proc. 1976 IEEE Conf. on Decision and Control, Dec. 1976, pp. 398-403.
3. Tou, J. T., Kao, D. B., and Chang, Y. S., "Pictorial texture analysis and synthesis", Proc. Third Int. Joint Conf. on Pattern Recognition, Coronado, Calif., 1976.
4. Franks, L. E., "A model for the random video process", Bell System Tech. Jnl., Vol. 45, April 1966, pp. 609-630.
5. Habibi, A., "Two dimensional Bayesian estimate of images", Proc. IEEE, Vol. 60, July 1972, pp. 878-883.
6. Huang, T. S., "The subjective effect of two-dimensional pictorial noise", IEEE Trans. on Inf. Theory, Vol. IT-11, Jan. 1965, pp. 43-53.
7. Jain, A. K. and Angel, Edward, "Image restoration, modelling, and reduction of dimensionality", IEEE Trans. on Computers, Vol. C-23, May 1974, pp. 470-476.
8. Kretzmer, E. R., "Statistics of television signals", Bell System Tech. Jnl., Vol. 31, July 1952, pp. 751-763.
9. Longuet-Higgins, M. S., "The statistical analysis of a random moving surface", Phil. Trans. Roy. Soc. London, Vol. A249, Feb. 1957, pp.321-387.
10. Longuet-Higgins, M. S., "Statistical properties of an isotropic random surface", Phil. Trans. Roy. Soc. London, Vol. A250, Oct. 1957, pp. 151-171.

11. Longuet-Higgins, M. S., "On the statistical distribution of the heights of sea waves", J. Mar. Res., Vol. 11, pp. 245-266.

12. McCormick, B. H. and Jayaramamurthy, S. N., "Time series model for texture synthesis", Int. J. Computer and Information Sciences, Vol. 3, 1974, pp. 329-343.

13. Pierson, W. J., "A unified mathematical theory for the analysis, propagation and refraction of storm generated surface waves", Dept of Meteorology, N. Y. University.

14. Panda, D. P., "Statistical analysis of some edge operators", University of Maryland Computer Science T. R. 558, July 1977.

15. Panda, D. P., "Statistical properties of thresholded images", University of Maryland Computer Science T. R. 559, July 1977.

16. Hunt, B. R., "Bayesian methods in nonlinear digital image restoration", IEEE Trans. on Computers, Vol. C-26, Mar. 1977, pp. 219-229.

17. Hunt, B. R. and Cannon, T. M. "Nonstationary assumptions for Gaussian models of images", IEEE Trans. on Syst., Man., and Cyb., Vol. SMC-6, Dec. 1976, pp. 876-882.

18. Nahi, N. E. and Jahanshahi, M. H., "Image boundary estimation", IEEE Trans. on Computers, Vol. C-26, Aug. 1977, pp. 772-781.

19. Schachter, B. and Ahuja, N., "A survey of random pattern generation processes", University of Maryland Computer Science T. R. 549, July 1977.

20. Rosenfeld, A. and Kak, A. C., Digital Picture Processing, Academic Press, New York, 1976.

21. Whittle, P., "On stationary processes in the plane", Biometrika, Vol. 41, 1954, pp. 434-449.

22. Wong, E., "Two dimensional random fields and representation of images", SIAM Jnl. Appl. Math., Vol. 16(4), 1968, pp. 756-770.

x	x	x	X	X	X	X
x	x	x	X	X	X	X
x	x	x	X	X	X	X
x	x	x	O	X	X	X
x	x	x	x	X	X	X
x	x	x	x	X	X	X
x	x	x	x	X	X	X

Fig. 1

		j-2	j-1	j
		⋮	⋮	⋮
i-2	...	X	X	X
i-1	...	X	X	X
i	...	X	X	O

Fig. 2

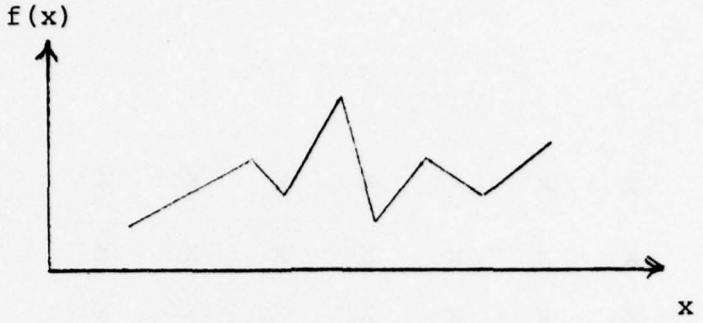
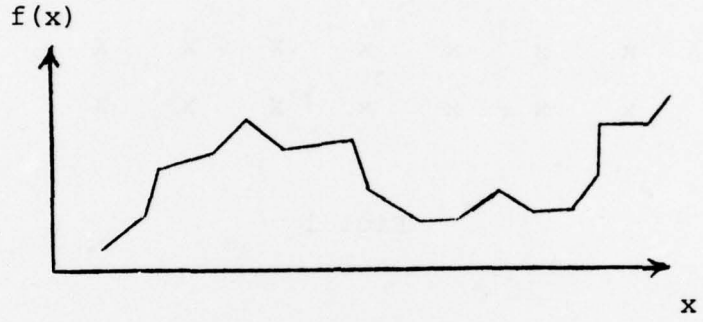
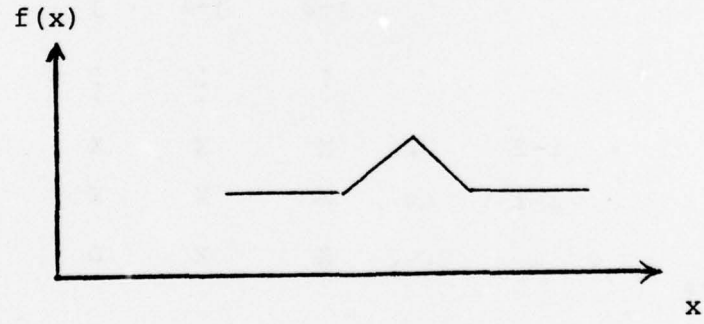


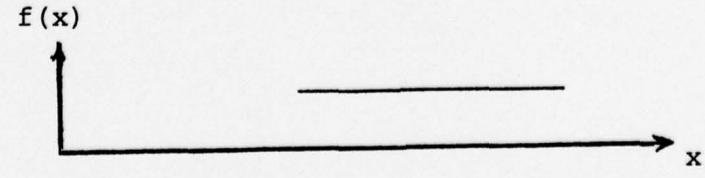
Fig. 3



(a)

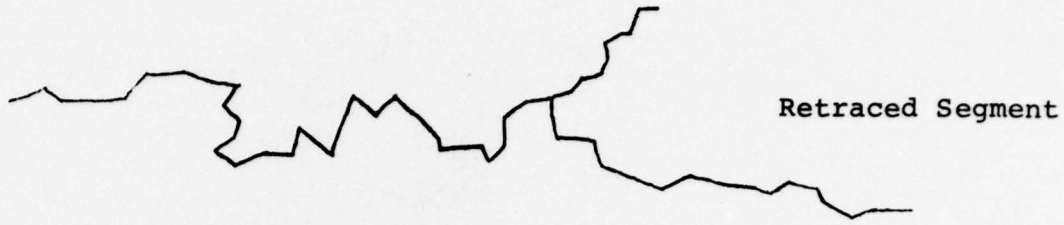


(b)

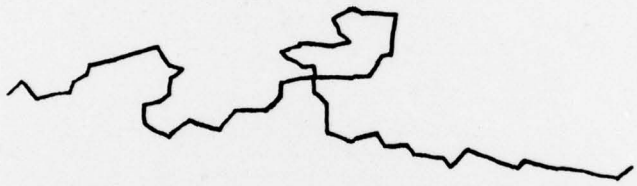


(c)

Fig. 4



(a)



(b)

Fig. 5

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFOSR-TR-78-0127	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): MOSAIC MODELS FOR IMAGE ANALYSIS AND SYNTHESIS	5. TYPE OF REPORT & PERIOD COVERED: Interim rept.
	6. PERFORMING ORG. REPORT NUMBER: TR-607
7. AUTHOR(s): Narendra Ahuja	8. CONTRACT OR GRANT NUMBER(s): AFOSR-77-3271
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University of Maryland, Computer Science Center, College Park, MD 20742	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A2
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, DC 20332	12. REPORT DATE: November 1977
	13. NUMBER OF PAGES: 46
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

image processing
pattern recognition
image modelling
stochastic processes
random geometry

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report is concerned with a class of models for natural images that treat an image as a mosaic, composed of patches generated by a random geometric process. These models are contrasted with conventional statistical image models. Proposed applications of such models in image analysis and synthesis are discussed.

403078

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)