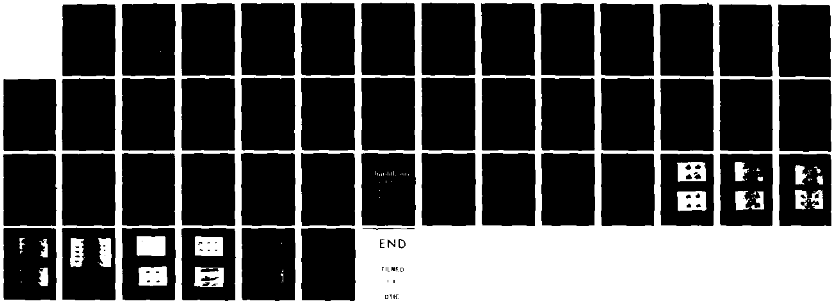
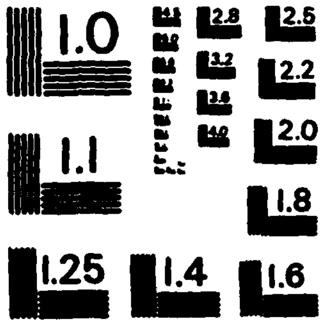
AD-A121 418 MONOTONE EQUIVARIANT SEGMENTATION TECHNIQUES(U)
MASSACHUSETTS UNIV AMHERST DEPT OF MATHEMATICS AND
STATISTICS M F JANOWITZ SEP 82 TR-J8201

UNCLASSIFIED N00014-79-C-0629 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 18201	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Monotone equivariant segmentation techniques		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) M. F. Janowitz		8. CONTRACT OR GRANT NUMBER(s) N00014-79-
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Massachusetts Amherst, MA 01003		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 121405
11. CONTROLLING OFFICE NAME AND ADDRESS Procuring Contract Officer Office of Naval Research Arlington, VA 22217		12. REPORT DATE Sept., 1982
		13. NUMBER OF PAGES 46
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Office of Naval Research, Resident Represent. Harvard University, Gordon McKay Laboratory Cambridge, MA 02138		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)

DTIC ELECTE NOV 9 1982 S D B

18. SUPPLEMENTARY NOTES

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

Image segmentation, thrshholding, cluster analysis, motone equivariance

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

An earlier model for nonhierarchical clustering is extended so that it encompases the ordinal image segmentation problem. A theoretical discussion within the model is provided for certain segmentation techniques. The discussion is also carried out within the framework of a Bernoulli process. The techniques are illustrated on real and simulated data.

82 3

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

MONOTONE EQUIVARIANT SEGMENTATION TECHNIQUES

by

M. F. Janowitz

Department of Mathematics and Statistics
University of Massachusetts
Amherst, MA 01003

1. Introduction. An order theoretic model for hierarchical cluster analysis was introduced in [5] and further developed in [6], [7], and [8]. The model was broad enough to include earlier models due to Jardine and Sibson [11] and Matula [12]. The purpose of the present paper is to show that the model also includes ordinal non-hierarchical clustering, and in particular may be viewed as a model for ordinal image segmentation. Section 2 contains some background material from the thoery of partially ordered sets, while section 3 relates this material to the problems of image segmentaion. The material is approached via a simple statistical model in section 4, and section 5 relates the present model to certain other segmentation techniques. The final section of the paper outlines how the theory of section 3 may be used to construct a segmentation program, and illustrates the operation of this program on real data.

2. Background material from the theory of partially ordered sets.

A basic familiarity with the theory of partially ordered sets and lattices will be assumed. Despite this, it will be convenient to specifically develop certain notions here. Unless otherwise specified, all partial orders will be denoted $\leq$, with $\subseteq$ reserved for set inclusion. The symbol $P(X)$ will represent the Boolean algebra formed by the subsets of the set X, when ordered by set inclusion. Unions and intersections will be represented as usual by the symbols $\cup$ and $\cap$, and it turns out to be convenient to let R denote the nonnegative real numbers ordered in the usual manner.

Let P,Q be partially ordered sets. A mapping $\phi: P \to Q$ is called isotone if $p_1 \leq p_2$ in P implies that $\phi(p_1) \leq \phi(p_2)$ in Q. It is said to be residuated in case it is isotone and there is an isotone mapping $\phi^*: Q \to P$ such that $p \leq \phi^*\phi(p)$ and $q \geq \phi\phi^*(q)$ hold for every $p \in P$, $q \in Q$. This mapping ϕ^* is called the residual mapping associated with ϕ, and is uniquely determined by ϕ. Likewise, ϕ is completely determined by ϕ^*, and for a given residual mapping θ, it will often be convenient to let θ^* denote the residuated mapping with which θ is associated. Let $\mathrm{Res}(P,Q)$ denote the set of all residuated mappings of P into Q, writing $\mathrm{Res}(P)$ in place of $\mathrm{Res}(P,P)$: dually, let $\mathrm{Res}^*(Q,P)$ denote the set of all residual mappings of Q into P, with $\mathrm{Res}^*(Q)$ having its expected meaning. The reason that residuated and residual mappings crop up is very simple. It was shown in [5], Lemma 4.1,

p. 60 that a residual mapping C from R into $P(X)$, where X is a finite nonempty set, may be characterized by the following three conditions:

(1) C is isotone

(2) There is an element h of R such that $C(h) = X$.

(3) Given $h \in R$, there corresponds a positive real number δ such that $C(h) = C(h+\delta)$.

If X happens to be the set of all 2 element subsets of the set P, then this says that C being residual is equivalent to C being a <u>numerically stratified clustering</u> in the sense of Jardine and Sibson [11], p. 61.

The theory of residuated mappings is rather extensively developed in [3], and the reader should refer to that source, should further information be needed.

3. <u>The image segmentation problem</u>. The underlying input data may be thought of as a function F defined on a bounded closed rectangle and taking values either in R or in R^p for some positive integer p. The values of F might represent brightness, temperature, or some other attribute of the rectangle. The function F is transformed by a sensing device into a function $\bar{F}$ having the same range as F, but whose domain is $X = \{1,2,\ldots,m\} \times \{1,2,\ldots,n\}$ for suitable positive integers m and n. The idea here is to divide the original rectangle into

finitely many subregions by means of a rectangular grid, and on each such subregion, let $\bar{F}$ somehow summarize the values of F. The actual input data that is presented to the observer then is not the original function F, but rather the summary function $\bar{F}$ defined on X; alternately, the input data may be thought of as an m by n matrix taking values in R or R^n. Unfortunately, the process of going from F to $\bar{F}$ has associated with it some noise, so $\bar{F}$ represents a distorted view of the actual data. The idea of the segmentation process is to recapture the principal regions represented by $\bar{F}$. A more precise and detailed description of the assumptions relating to this are contained in [9] and [10]. For that reason they will not be repeated here. We should mention though that our attention will be restricted to the case where F takes values in R, with the multifeature situation to be investigated in a later paper.

Let us see how to embed the segmentation problem into the model for hierarchical clustering that was presented in [5]. Think of the input data either as a mapping $D:X \to R$, where $X = \{1,2,\dots,m\} \times \{1,2,\dots,n\}$, or as an m by n matrix A with entries from R. The idea is to somehow represent the information provided by A as a matrix B (possibly of a size smaller than A) that has fewer distinct levels of entries. To properly place this problem in an order theoretic setting, recall ([5], p. 61) that D extends naturally to a residuated mapping $D^*: P(X) \to R$ by letting

$D^*(M) = \vee \{D(m): m \in M\}$ for all subsets M of X. If D^+ denotes the unique residual mapping associated with D^*, we have established a bijection between mappings from X into R and residual mappings from R into $P(X)$. Thus the image segmentation problem may be viewed as the study of transformations of $Res^*(R,P(X))$ into $Res^*(R,P(Y))$ where $Y \subseteq X$. Transformations such as this will be called <u>segmentation methods</u>. This places the segmentation problem squarely within the framework of the model described in [5].

For input data having only ordinal significance, we would like to describe the class of segmentation methods that should properly be used. In order to do this, we need the notion of compatibility ([5], p. 68). Let F be a segmentation method and θ a residual mapping on R. To say that F is θ-compatible is to say that for every $C \in Res^*(R,P(X))$, $F(C \circ \theta) = F(C) \circ \theta$. For data having only ordinal significance, one should insist on θ-compatibility for all order automorphisms of R and this leads to a class of segmentation techniques that are called <u>monotone equivariant</u>. These are characterized ([7], Theorem 1, p. 149) as follows: Let $C \in Res^*(R,P(X))$. To say that h $(h \in R)$ is a <u>splitting level</u> of C is to say that h is the image of some subset of X under the residuated mapping C^* associated with C. If $0 < h_1 < \ldots < h_t$ denote the splitting levels of C, then to say that the segmentation method F is monotone equivariant is to say that:

(i) every splitting level of $F(C)$ is a splitting level of C,

(ii) the sequence $FC(0) \leq FC(h_1) \leq \ldots \leq FC(h_t)$ depends only upon the sequence $C(0) < C(h_1) < \ldots < C(h_t)$ and is independent of the actual values of the h_i's.

(iii) conditions (i) and (ii) hold for every $C \in Res^*(R,P(X))$.

One can associate with each subset T of $Res^*(R)$ the collection $\alpha(T)$ of all segmentation methods that are θ-compatible for every $\theta \in T$; conversely, for any set S of segmentation methods, one can associate the set $\beta(S)$ of all $\theta \in Res(R)$ such that F is θ-compatible for every F in S. Naturally, for this to be meaningful, one must think of a segmentation method as a mapping $F: Res^*(R,P(X)) \to Res^*(R,P(Y))$ where Y is a fixed subset of X. In any event, the mappings (α,β) define a galois correspondence in the sense of [2], p. 124 between the residual mappings on R and the segmentation methods on X. It generally turns out that the galois closed objects (i.e., those objects that appear as images of α or β) of such correspondences have some significance. We have already seen that the monotone equivariant cluster methods appear as the image under α of the set of order automorphisms of R, so they are galois closed. F. Baulieu [1] has among many other things characterized all closed classes that are contained in this one. Here is a description of these classes:

(S1) Flat methods. $FC(h)$ depends only upon $C(h)$.

(S2) Semiflat methods. $FC(h)$ depends upon $C(h)$ and $C(0)$.

(S3) Divisive methods. $FC(h_i)$ depends upon $C(h_i)$, $C(h_{i+1}),\ldots,C(h_t)$ where h_t is the highest splitting level of C.

(S4) SF(1) methods. $FC(0)$ depends only upon $C(0)$; for $h_i \neq 0$, $FC(h_i)$ depends on $C(h_i)$, $C(h_{i+1}),\ldots,C(h_t)$.

(S5) Monotone equivariant methods.

The related sets of residual mappings, (Ti) corresponding to (Si), are:

(T1) Res(R)

(T2) The set of all residual mappings θ on R for which $\theta(0) = 0$.

(T3) Those residual maps whose range is of the form $\{h: h \geq k\}$ for some fixed element k of R; i.e., whose range is a principal filter of R.

(T4) Those residual maps whose range is the union of $\{0\}$ with a principal filter of R.

(T5) The set of all injective residual mappings on R.

The simplest of the closed classes of segmentation methods is of course the flat methods, and it is to this class that we now direct our attention. Needless to say, such methods are easy to describe and easy to implement. We are given a fixed finite nonempty set X and an increasing sequence

$$M_1 \subset M_2 \subset \ldots \subset M_t = X$$

of subsets of X, where $M_i = C(h_i)$, and wish to produce a sequence

$$N_1 \subseteq N_2 \subseteq \dots \subseteq N_t = X$$

of subsets of X that somehow summarize or better represent the underlying picture than does the input data to the problem. In that N_i depends only upon M_i, this amounts to defining an isotone mapping ψ on $\mathcal{P}(X)$.

Having recognized the need to define an isotone mapping on $\mathcal{P}(X)$, it is appropriate to decide upon reasonable candidates for these mappings. The idea is to attach some spatial significance to the decision as to whether a given point belongs to the output of the mapping. In other words, the decision as to whether x belongs to $\psi(M)$ should be based upon all points in some small region surrounding x. This may be precisely stated by saying that the mapping ψ shall be <u>point-based</u> in that for each $x \in \psi(X)$ there is a subset $N(x)$ of X containing x such that:

(PB1) $\psi(\emptyset) = \emptyset$ and $\psi(N(x)) \neq \emptyset$.

(PB2) For $A \subseteq N(x)$, if $\psi(A) \neq \emptyset$, then $x \in \psi(A)$.

(PB3) For $M \subseteq X$, $x \in \psi(M)$ if and only if $x \in \psi(M \cap N(x))$.

It is immediate that for $M \subseteq X$, $\psi(M) = \bigcup_{x \in \psi(X)} \psi(M \cap N(x))$. So that the output will be independent of the polarity of the input image, it sometimes turns out to be useful to insist that ψ <u>preserve complements</u> in that

(PB4) For $M \subseteq X$, $\psi(X \setminus M) = \psi(X) \setminus \psi(M)$.

The local version of (PB4) is the content of Theorem 1.

Theorem 1. *For a point-based isotone mapping γ on $P(X)$, axiom (PB4) is equivalent to the assertion that for each subset A of $N(x)$, exactly one of $\gamma(A)$ and $\gamma(N(x)\setminus A)$ will be nonempty.*

Proof: Let γ satisfy (PB4). If $x \in \gamma(A)\cap\gamma(N(x)\setminus A)$, then $x \in \gamma(A) \cap \gamma(X\setminus A)$, contrary to (PB4). For the converse implication, assume that γ is point-based and that for $A \subseteq N(x)$, exactly one of $\gamma(A)$ and $\gamma(N(x)\setminus A)$ is nonempty. If now $y \in \gamma(X\setminus M)$, then $y \in \gamma((X\setminus M)\cap N(y))$, so $y \notin \gamma(M\cap N(y))$, and consequently, $y \notin \gamma(M)$. Conversely, $y \notin \gamma(M) \Rightarrow y \notin \gamma(M\cap N(y))$, so $y \in \gamma(N(x)\setminus M) \subseteq \gamma(X\setminus M)$.

Now let γ be point-based with an associated family of neighborhoods $N(x)$, $x \in \gamma(X)$. If $M(x)$ is a second such family of neighborhoods, it is clear that so also is $M(x) \cap N(x)$. In view of this, there is no harm in assuming

(PB5) If $M(x)$ satisfies (PB1) through (PB3), then $N(x) \subseteq M(x)$.

Such a minimal family of subsets of X will be called the *system of neighborhoods associated with* γ. Unless otherwise specified, when we speak of a point-based isotone mapping γ, it will always be assumed that the family $\{N(x):x \in \gamma(X)\}$ represents this system of neighborhoods.

<u>Definition</u>. The point-based isotone mapping φ is said to be <u>frequency-defined</u> if there exist positive integers j and k such that for each $x \in \varphi(X)$,

(i) $\#N(x) = k$, and

(ii) $x \in \varphi(M)$ for $M \subseteq X$ if and only if $j \leq \#(M \cap N(x))$.

Here $\#A$ denotes the cardinality of the set A.

<u>Theorem 2. Let φ be a point-based, complement-preserving isotone mapping on $P(X)$. Suppose that the regions $N(x)$ all have the same cardinality, and that for each $x \in \varphi(X)$, A_x has minimal cardinality among those subsets A of $N(x)$ for which $x \in \varphi(A)$. Necessary and sufficient conditions for φ to be frequency-based are that $k = \#N(x)$ be odd, and that $\#A_x = (1+k)/2$ for all $x \in \varphi(X)$.</u>

<u>Proof</u>: Let φ be frequency-based. If $\#A_x \leq k/2$, then $\#(N(x) \setminus A_x) \geq k/2$, and there is a subset B of $N(x) \setminus A_x$ such that $\#B = \#A_x$. Since $\varphi(B) = \emptyset$, this produces a contradiction. If k is even, taking $B \subseteq N(x)$ with cardinality $k/2$ produces a similar contradiction. Thus k is odd and $\#A_x > k/2$. If $\#A_x > (1+k)/2$, we may take $B \subseteq A_x$ so that $\#B = (1+k)/2$. Since $\#(N(x) \setminus B) < \#B$, this tells us that $\varphi(B) = \emptyset = \varphi(N(x) \setminus B)$, contrary to (PB4). Thus $\#A_x$ must equal $(1+k)/2$. Suppose conversely that $k = \#N(x)$ is odd and that $\#A_x$ is necessarily $(1+k)/2$. We are to show that φ

is frequency-based. By construction of A_x, $x \in \gamma(B)$ for $B \subseteq N(x)$ implies that $\#B \geq \#A_x = (1+k)/2$. If on the other hand, $\#B < (1+k)/2$, then $\#(N(x)\setminus B) \geq (1+k)/2$, so $\gamma(N(x)\setminus B) \ni x$, and consequently, $x \notin \gamma(B)$.

One is often interested in the extent to which a segmentation algorithm blurs an image. The merging of clusters that should not be merged is one of the ways in which such blurring occurs. To illustrate this concretely, let us consider a frequency-defined mapping γ that operates on 3 by 3 neighborhoods centered on points. Suppose X has 5 rows and 5 columns, and that M, N are the regions denoted by 1's, 2's in Fig. 2(a). In Fig. 2 (b), the members of $\gamma(M)$, $\gamma(N)$ are denoted by 1's, 2's, while in Fig. 2(c), the members of $\gamma(M \cup N)$ are indicated by 3's. All remaining entries are 0. Notice that $\gamma(M \cup N) \supsetneq \gamma(M) \cup \gamma(N)$. In connection with this, we shall say

(a)						(b)						(c)				
1	1	1	1	1		0	0	0	0	0		0	0	0	0	0
1	1	1	1	1		0	1	1	1	0		0	3	3	3	0
0	0	0	0	0		0	0	0	0	0		0	3	3	3	0
2	2	2	2	2		0	2	2	2	0		0	3	3	3	0
2	2	2	2	2		0	0	0	0	0		0	0	0	0	0

Fig. 2 Illustration of blurring. See text for explantation.

that M, N are γ-separated for the isotone point-based mapping γ in case there is no set of the form $N(x)$ that intersects both M and N. We then have

Theorem 3. Let γ be a point-based isotone mapping on $P(X)$. If M, N are γ-separated, then $\gamma(M \cup N) = \gamma(M) \cup \gamma(N)$.

Proof: Evidently, we need only show that $\gamma(M \cup N) \subseteq \gamma(M) \cup \gamma(N)$. Accordingly, let $x \in \gamma(M \cup N)$, so $(M \cup N) \cap N(x) \neq \emptyset$. Since M, N are γ-separated, we must have $(M \cup N) \cap N(x) = M \cap N(x)$ or $N \cap N(x)$, whence $x \in \gamma(M) \cup \gamma(N)$.

Along these lines, the mapping γ will be called a join homomorphism in case $\gamma(M \cup N) = \gamma(M) \cup \gamma(N)$ for all subsets M, N of X. There is a dual notion of meet homomorphism, and to say that γ is a homomorphism is to say that it is both a join and a meet homomorphism. The next theorem shows that these conditions are extremely powerful, and will only be met in trivial situations.

Theorem 4. Let γ be a point-based isotone mapping on $P(X)$. Then (1a)<=>(1b), (2a)<=>(2b) and (3a)<=>(3b).

(1a) γ is a join homomorphism.

(1b) For each $x \in \gamma(X)$ there is a subset A_x of $N(x)$ such that $x \in \gamma(M)$ if any only if $M \cap A_x \neq \emptyset$.

(2a) γ is a meet homomorphism.

(2b) Corresponding to each $x \in \gamma(X)$, there is a subset B_x of $N(x)$ such that $x \in \gamma(M)$ if and only if $B_x \subseteq M$

(3a) γ is a homomorphism.

(3b) <u>For each</u> $x \in \gamma(X)$ <u>there corresponds an element</u> v_x <u>of</u> $N(x)$ <u>such that</u> $x \in \gamma(M)$ <u>if and only if</u> $v_x \in M$.

<u>Proof</u>: (1a) $\Longrightarrow$ (1b). If γ is a join homomorphism, let C be the union of all subsets of $N(x)$ that are mapped to $\emptyset$ by γ, and note that $\gamma(C) = \emptyset$. Observing that $\gamma(M \cap N(x)) = \emptyset$ if and only if $M \cap N(x) \subseteq C$, we may now take $A_x = N(x) \setminus C$.

(1b) $\Longrightarrow$ (1a) If $x \in \gamma(M \cup N)$, then $(M \cup N) \cap A_x \neq \emptyset$, so $M \cap A_x$ or $N \cap A_x \neq \emptyset$. Hence $x \in \gamma(M) \cup \gamma(N)$.

(2a) $\Longrightarrow$ (2b) If B_x is the intersection of all subsets C of $N(x)$ for which $\gamma(C) \neq \emptyset$, then $x \in$ every $\gamma(C)$, so $x \in \gamma(B_x)$. (2b) is now clear.

(2b) $\Longrightarrow$ (2a) is obvious, as is (3b) $\Longrightarrow$ (3a).

(3a) $\Longrightarrow$ (3b) If B_x is defined as in (2), then $\gamma(B_x) \neq \emptyset$, but $\gamma(C) = \emptyset$ for every proper subset C of B_x. It is immediate that $B_x = \{v_x\}$, and (3b) now follows.

<u>Remark</u> 5. Noting that if γ in the theorem is complement-preserving, then it is a join homomorphism if and only if it is a meet homomorphism, it follows that for such mappings the six conditions of the theorem are mutually equivalent.

If $M \subseteq X$ is acted upon by a rigid motion to produce N, one would naturally want that same rigid motion acting upon $\gamma(M)$ to produce $\gamma(N)$. A careful formulation of this idea shows that this is too much to hope for. To see why, we need the concept of a translation on X: this is a mapping of the form $T_{p,q}$ where p,q are fixed integers and $T_{p,q}(i,j) = (i+p, j+q)$. Unless $p = 0 = q$, the domain and range of $T_{p,q}$ will be proper subsets of X. We agree to call the point-based isotone mapping translation-invariant provided it satisfies:

(PB6) If $N(x)$ is contained in the domain of the translation T, and if $y = T(x)$, then $N(y) = T(N(x))$, and for $A \subseteq N(x)$, $x \in \gamma(A)$ if and only if $y \in \gamma(T(A))$.

Remark 6. If γ is a point-based, translation-invariant isotone mapping on $P(X)$, and if $M \subseteq$ domain (T) with $M \cup T(M) \subseteq \gamma(X)$, it is true that $T(\gamma(M)) = \gamma(T(M))$. This may fail, however, if either M or $T(M)$ is not contained in $\gamma(X)$. This is illustrated in Fig. 3. Here $X = S \times S$, where $S = \{1,2,3,4,5\}$. The mapping γ is defined by looking at 3 by 3 neighborhoods centered on the points of X, and saying that for $A \subseteq N(x)$, $x \in \gamma(A)$ iff $\#A \geq 5$. This choice of γ satisfies (PB1) through (PB6), and is even frequency-defined. Let $x = (2,2)$ and $T = T_{1,1}$, so that $T(x) = (3,3)$. Then Fig. 3(a) shows $N(x)$, Fig. 3(b) shows $T(N(x)) = N(T(x))$, while $\gamma(N(x))$ and $\gamma(T(N(x)))$ are displayed in (c) and (d). The reason that $T(\gamma(N(x))) \neq \gamma(T(N(x)))$ is that $T(N(x))$ is contained in $\gamma(X)$, while $N(x)$ is not.

1 1 1 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
1 1 1 0 0	0 1 1 1 0	0 1 1 0 0	0 0 1 0 0
1 1 1 0 0	0 1 1 1 0	0 1 0 0 0	0 1 1 1 0
0 0 0 0 0	0 1 1 1 0	0 0 0 0 0	0 0 1 0 0
0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
(a) $M = N(x)$	(b) $T(M)$	(c) $\gamma(M)$	(d) $\gamma(T(M))$

Fig. 3 See text for explanation.

<u>Remark</u> 7. For translation-invariant mappings γ on $P(X)$, the conditions expressed in Theorems 2,3 and 4 can be simplified, as one need only state them for a single member of $\gamma(X)$. The details of this are left to the reader. We also mention without proof that versions of the above results are true for atomistic orthomodular lattices([2], p. 53).

<u>Remark</u> 8. The actual construction of a point-based isotone mapping on $P(X)$ can now be easily understood. One chooses a system of neighborhoods $N(x)$ for points x in some subset Y of X, and defines a family $(\gamma_x)_{x \cdot Y}$, where γ_x is an isotone mapping on $N(x)$ such that :

(1) $\gamma_x(\phi) = \phi$, and $\gamma_x(N(x)) = \{x\}$.

If one wants to produce a complement-preserving mapping, then one also wants

(2) If $A \subseteq N(x)$, then exactly one of $\gamma_x(A)$ and $\gamma_x(N(x) \setminus A)$ is non-empty.

The mapping γ is now defined by the rule $x \in \gamma(M)$ if and only if $\gamma_x(M \cap N(x)) = \{x\}$. This is the technique that will be used for the remainder of the paper.

Remark 9. Though it might not necessarily make sense to use point-based mappings that are not frequency-defined, it is easy to construct examples of them. To see this, let $N(x)$ be a 3 by 3 region centered on x. For $A \subseteq N(x)$, let $\gamma_x(A) = \{x\}$ if at least 2 out of the 3 rows of $N(x)$ each has at least 2 members of A; otherwise, let $\gamma_x(A) = \phi$. Do this for each x that is not in an outer row or column of X, and use the construction of Remark 8 to define a mapping γ on $P(X)$. The result is a point-based, complement-preserving isotone mapping that is not frequency-defined. Clearly this example is representative of an entire class of such mappings. Other candidates for γ are pattern-defined mappings. These put x in $\gamma_x(A)$ if A contains a subset of a desired type. For 3 by 3 neighborhoods, such desired types might include sets such as

$$\begin{matrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{matrix} \qquad \text{or} \qquad \begin{matrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{matrix} \; .$$

4. Underlying statistical considerations. We have just seen that the construction of a flat segmentation method involves the

definition of an isotone mapping on $P(X)$. If the input data were correct, then the most reasonable choice for such a mapping might be the identity map or a map of the form $M \to M \cup Y$ for some subset Y of X. The idea though is to use the isotone mapping γ to produce an estimate of the true data that is in some sense better than the one provided by the input data. Specifically, let us assume that the data represents a subset M^* of X, but that the input data produced M instead. One wants to estimate M^* by $\gamma(M)$. As in the last section, it is assumed that $X = \{1,2,\ldots,m\} \times \{1,2,\ldots,n\}$. Let $\{N(x)\}$ be a system of neighborhoods of points x in some subset Y of X. To say that a point (i,j) is an interior point of the set M^* will be to say that for $x = (i,j)$, $N(x) \subseteq M^*$; a similar definition applies to the complement of M^*. If $N(x)$ intersects both M^* and its complement, then x will be called a <u>boundary point</u> of M^*. Thus X is divided into 4 regions: (1) the interior of M^*, (2) the interior of $X \backslash M^*$, (3) the boundary of M^*, and (4) those points which do not have neighborhoods (i.e., $X \backslash Y$). This is concretely illustrated in Fig. 4. Here we take $m = 10$, $n = 7$, and $N(x)$ a 3 by 3 region centered on x. Y would then consist of all points not in an outer row or column of X. The members of M^* are denoted by 1's in Fig. 4(a) with those of $X \backslash M^*$ represented by 0's. The resulting 4 regions are displayed in Fig. 4(b) and are numbered as above.

0	0	0	0	0	0	0
0	0	0	0	0	1	1
0	0	0	0	1	1	1
0	0	0	1	1	1	1
0	0	1	1	1	1	1
0	0	1	1	1	1	1
0	0	1	1	1	1	1
0	0	0	1	0	1	1
0	0	0	0	1	1	1
0	0	0	0	0	1	1

(a) Original picture

4	4	4	4	4	4	4
4	2	2	3	3	3	4
4	2	3	3	3	1	4
4	3	3	3	3	1	4
4	3	3	3	1	1	4
4	3	3	1	1	1	4
4	3	3	3	3	3	4
4	3	3	3	3	3	.
4	2	3	3	3	3	4
4	4	4	4	4	4	4

(b) 3 by 3 neighborhoods

Fig. 4. See text for explanation.

A crude model may be constructed by assuming that membership in M has probability p $(p > 0.5)$ of providing a correct estimate of membership in M^*, and that membership in $X \backslash M$ has a probability $q(q > 0.5)$ of providing a correct estimate for membership in M. Let us also assume that these probabilities are each independent, bearing in mind that in practice this is sometimes not so. Realistically, one should also notice that boundary points will have a lower probability of correct classification than do interior points, but this is an issue that will be temporarily ignored. The gain of an isotone mapping γ is defined to be the sum of the probability of correct classification by γ for an interior point of M^* and an interior point of $X \backslash M^*$, and is denoted $G(\gamma)$. Until further notice, it will be assumed that γ is a point-based, translation-invariant isotone mapping on $P(X)$. It will prove convenient for a fixed $N(x)$ to let $T_\gamma = \{A\colon A \subseteq N(x),\ x \in \gamma(A)\}$ and $F_\gamma = \{A\colon A \subseteq N(x),\ \gamma(A) = \Phi\}$.

Letting A^i denote a typical subset of $N(x)$ having cardinality i, we have then have

$$G(\gamma) = \Sigma\{p^i(1-p)^{k-i} : A^i \in T_\gamma\} + \Sigma\{(1-q)^i q^{k-i} : A^i \in F_\gamma\},$$

where $k = \#N(x)$. In order to see the role of frequency-defined mappings in this model, we present

Lemma 10. *Let* γ *be complement-preserving with* $k = \#N(x)$ *odd*, and let γ' *be the frequency-defined, point-based isotone mapping having the same neighborhood system as* γ. *Then* $G(\gamma) \le G(\gamma')$.

Proof: Since γ is complement-preserving, $\#T_\gamma = \#F_\gamma = 2^{k-1}$. The same is true for γ'. Thus a bijection ϕ may be defined on $P(N(x))$ such that ϕ maps T_γ onto $T_{\gamma'}$, and F_γ onto $F_{\gamma'}$. We may even take ϕ to be the identity map on $(T_\gamma \cap T_{\gamma'}) \cup (F_\gamma \cap F_{\gamma'})$. For $A \in T_\gamma \cap F_{\gamma'}$, $\#A < j = (1+k)/2$, while $\#\phi(A) \ge j$, so $\#A < \#\phi(A)$. A similar argument shows that for $A \in F_\gamma \cap T_{\gamma'}$, $\#A > \#\phi(A)$. Using ϕ to match the terms of $G(y)$ with those of with those of $G(\gamma')$, it is clear that each term of $G(\gamma)$ is no larger than the corresponding term of $G(\gamma')$, so $G(\gamma) < G(\gamma')$.

Theorem 11. *If* $p = q$, *and if* γ *and* γ' *have the same system* $N(x)$ *of neighborhoods with* γ' *frequency-defined, then* $G(\gamma) \le G(\gamma')$.

<u>Proof</u>: In view of Lemma 10, we need only produce a complement-preserving point-based isotone mapping γ'' with $N(x)$ as neighborhoods such that $G(\gamma) < G(\gamma'')$. This will be accomplished by making use of Remark 8. Choose some fixed $N(x)$, and consider subsets A of $N(x)$. If $x \in \gamma(A) \cap \gamma(N(x)\setminus A)$ with $\#A < j$, take $\gamma''(A) = \{x\}$ and $\gamma''(N(x)\setminus A) = \phi$. If $\gamma(A) = \phi = \gamma(N(x)\setminus A)$ with $\#A < j$, do the same. Otherwise, let $\gamma''(A) = \{x\}$ if $x \in \gamma(A)$ and $\gamma''(A) = \phi$ if $\gamma(A) = \phi$. We need to show that γ'' is isotone. The proof will be by contradiction. Suppose then that $A \subset B \subseteq N(x)$ with $\gamma''(A) \not\subseteq \gamma''(B)$. Then $\gamma''(A) = \{x\}$ and $\gamma''(B) = \phi$. By the construction of γ'', we must either have $x \in \gamma(A)$ or $\#A \geq j$.

<u>Case</u> 1. $x \in \gamma(A)$. Then also $x \in \gamma(B)$, so if $\gamma''(B) = \phi$, then $x \in \gamma(B) \cap \gamma(N(x)\setminus B)$ and $\#B < j$. Since $A \subset B$ implies $N(x)\setminus A \supset N(x)\setminus B$, this puts x in $\gamma(N(x)\setminus A)$, whence $\gamma''(A) = \phi$, a contradiction.

<u>Case</u> 2. $\#A \geq j$. Then also $\#B \geq j$. The only way for $\gamma''(B) = \phi$ is for $\gamma(B) = \phi$, so $\gamma(A) = \phi$ and consequently $\gamma''(A) = \phi$, again a contradiction.

This then establishes that γ'' is isotone. The construction of Remark 8 can now be used to extend γ'' to a point-based, complement-preserving isotone mapping on $P(X)$ having the same neighborhood system as γ. We must still show that $G(\gamma) < G(\gamma'')$. The changes in these sums occur from members of $(T_\gamma \cap F_{\gamma''}) \cup (T_{\gamma''} \cap F_\gamma)$. If $A^i \in T_\gamma \cap F_{\gamma''}$, then $i < j$, and the term corresponding to A^i in

$G(\gamma)$ is $p^i(1-p)^{k-i}$, while the corresponding term in $G(\gamma'')$ is $p^{k-i}(1-p)^i$. In that $i < k-i$ and $p > 0.5$, this represents an increase. Similarly, if $A^i \in T_{\gamma''} \cap F_{\gamma}$, then $i > j$, and again the term from $G(\gamma'')$ > the corresponding term from $G(\gamma)$.

When $p = q$, the above theorem shows that the "gain" from a flat segmentation method can be maximized by using a point-based isotone mapping that is complement-preserving and frequency-defined. This amounts to saying that such mappings maximize the probability of correct classification of interior points, as this probability is just half the gain. With this thought in mind, we shall direct our attention to mappings of the above type defined on k by k neighborhoods of points with k an odd integer. The j/k^2 rule with $j = (1+k^2)/2$ will be the unique such mapping on a k by k neighborhood, and the 3/5 rule will refer to the unique γ defined on the region consisting of x and the 4 points that are immediately to its North, South, East and West. These rules were introduced in a slightly different context in [9], and some discussion given to the probability of correct classification of points in the case of interior points as well as boundary points. These probabilities are all based on an underlying binomial distribution of $N(x)$, and will not be repeated here.

There is another item worth mentioning, however. To improve the classification of interior points, one simply enlarges the size of the k by k region. This is fine except that as k increases the number of boundary points also increases, and as was shown in [9], the probability of misclassification of boundary points can actually increase when the j/k^2 rule is used. This is illustrated by some simulations in Table 1. The segmentations are done on the data of Fig. 6(b) corrupted by various types of additive noise with the indicated signal to noise ratio. The theoretical probability of misclassification of interior points is compared to the actual probability based on 10 trials, and this is compared to the probability of misclassification in the entire picture. Note how the ratio of the probability for misclassification of interior points over that of the entire picture changes as the value of k increases from 5 to 9 to 25.

All of this suggests that there may be occasions when one wants to use flat segmentation method other than that provided by the j/k^2 rule. For 3 by 3 regions, one idea might be to define γ on $N(x)$ as follows: $\gamma(A) = \{x\}$ if $\#A \geq 6$ or if $\#A = 4$ or 5 and $x \in A$. Naturally this will increase the probability of misclassification of interior points, but may do a slightly better job near the boundary of a region. This is illustrated in Table 2, where this rule is compared to the $5/9$ rule for regions containing various mixtures of members of M and its complement. Notice the rather dramatic increase on probability of correct classification near the boundary as evidenced by the box relating to $\#M = 5$. A comparison of the two methods on real data is displayed in Fig. 5.

Noise Ratio	Rule	Actual image Observed	SD	Interior points Theoret.	Observed	SD	Ratio of Interior/actual
1.25	None	.2697	.0203		.2820	.0182	
	3/5	.1607	.0290	.1401	.1341	.0241	.83
	5/9	.1233	.0247	.0783	.0699	.0196	.57
	13/25	.0947	.0146	.0102	.0035	.0030	.037
1.00	None	.3100	.0199		.3178	.0262	
	3/5	.2218	.0336	.1874	.1935	.0377	.87
	5/9	.1784	.0271	.1221	.1256	.0276	.70
	13/25	.1381	.0272	.0283	.0333	.0217	.24
0.67	None	.3810	.0306		.3857	.0228	
	3/5	.3015	.0524	.2930	.2910	.0401	.97
	5/9	.2629	.0578	.2374	.2353	.0452	.90
	13/25	.2100	.0521	.1210	.1178	.0449	.56
0.50	None	.4170	.0265		.4168	.0163	
	3/5	.3564	.0447	.3469	.3361	.0333	.94
	5/9	.3276	.0611	.3026	.2902	.0340	.89
	13/25	.2739	.0669	.1987	.1810	.0391	.66
0.25	None	.4501	.0282		.4544	.0225	
	3/5	.4153	.0513	.4150	.4108	.0339	.99
	5/9	.3987	.0652	.3890	.3737	.0532	.94
	13/25	.3619	.0744	.3222	.3048	.0883	.84

Table 1. Simulated data based on 10 trials. Entries in the table represent probabilities of misclassification. The columns labelled SD refer to the standard deviation of this probability, and the "Ratio" column is simply the observed probability for interior points divided by the observed probability for the entire picture. The columns headed "Actual Image" refer to Fig. 6 with additive gaussian noise having the indicated signal to noise ratio. The remaining columns apply to the identical noise on the interior of a single region. The signal to noise ratio is the difference in grey levels that characterize the input set from its background divided by the standard deviation of the noise.

#M	#X-M	Prob. Correct	5/9 rule Mean	SD	Modified 5/9 rule Mean	SD
9	0	.9	99.80	0.51	99.55	0.74
		.8	97.55	1.72	96.30	2.26
		.7	90.50	3.12	86.65	4.61
		.6	73.30	5.04	70.65	4.27
8	1	.9	99.55	0.50	98.80	0.87
		.8	95.45	2.01	91.50	3.06
		.7	83.25	3.28	81.10	4.10
		.6	68.35	4.61	67.20	4.60
7	2	.9	98.10	1.48	96.20	1.60
		.8	89.50	2.62	86.90	3.18
		.7	77.25	3.48	76.20	3.19
		.6	64.45	4.82	62.45	3.20
6	3	.9	91.80	2.40	91.95	2.16
		.8	79.25	3.48	82.30	2.76
		.7	67.35	5.53	72.40	4.59
		.6	59.75	5.49	62.05	5.28
5	4	.9	72.55	3.94	88.25	2.95
		.8	62.25	4.30	77.55	3.83
		.7	57.70	5.01	65.40	5.30
		.6	55.70	5.12	58.75	4.23

Table 2. M is a subset of a 3 by 3 rectangle with the indicated cardinality. The Prob. Correct refers to the probability that the actual input data provides a correct estimate of membership in M. The entries in the columns labelled "Mean" refers to the mean percentage of probability of correct classification based on 20 trials, and the "SD" columns refer to the standard deviations of these means. See text for an explanation of the modified 5/9 rule.

Another approach to modifying a frequency-based mapping is to use some sort of pattern-based mapping. A way of implementing this is via the iterated 3 by 3 mean filter. This filter is based on a 5 by 5 neighborhood, and can either be viewed as a weighted filter on this neighborhood or as a 3 by 3 mean filter applied twice. (See [10] for a discussion of this filter.). Table 3 provides a comparison of the following 5 segmentation techniques: (1) original data; (2) 5/9 rule: (3) 5/9 rule but using iterated 3 by 3 mean for 4 or 5 members of M in the 3 by 3 neighborhood; (4) iterated 3 by 3 mean; (5) 13/25 rule. By Theorem 2, the 13/25 rule has the highest probability of correct classification for interior points, but as evidenced by Table 3, other segmentation techniques can give better overall performance. To see why this is so, one might notice that the 13/25 treats equally the following two data sets:

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix} \qquad\qquad \begin{matrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{matrix}$$

since they each have exactly 9 entries of 1. Yet the first data set is much more likely to reflect x being in the set M. On the other hand, the iterated 3 by 3 mean would produce values of 0.61 and 0.38 for these sets, thus putting x in M for the first but not the second data set.

Data	Error	Mean and SD of percentage incorrect classification: data	5/9 rule	modified 5/9 rule	iterated 3 by 3 mean	13/25 rule
A	U8	17.0	3.8	2.2	2.1	2.0
		1.1	0.6	0.6	0.6	0.5
	U15	33.6	17.9	12.8	[illegible]	10.1
		1.5	2.8	3.6	3.5	3.6
	U23	39.5	26.5	22.3	21.4	16.8
		2.4	5.0	6.2	6.3	5.5
B	U8	16.3	3.2	2.3	2.2	2.9
		1.7	0.8	0.5	0.6	0.7
	U15	32.7	16.5	10.6	9.8	8.3
		1.1	2.3	2.3	2.3	2.5
	U23	38.5	25.5	19.3	18.4	15.8
		2.2	4.3	4.6	4.6	4.6
C	U8	18.3	7.8	4.9	4.7	6.5
		1.7	1.5	1.2	1.2	1.3
	U15	33.1	21.3	16.1	15.4	16.0
		2.1	1.8	1.6	1.5	1.9
	U23	40.4	32.3	28.2	27.3	25.2
		2.9	7.2	7.2	7.3	6.0
D	U8	16.1	6.9	5.4	5.4	7.4
		2.1	1.8	1.6	1.5	1.9
	U15	32.0	18.6	14.0	13.5	15.4
		1.7	2.8	3.0	3.4	2.8
	U23	39.3	30.7	26.3	25.6	25.4
		2.4	4.8	5.1	5.2	6.1
A	G5	16.6	3.0	2.0	2.0	2.5
		1.3	0.4	0.4	0.4	0.5
	G10	32.2	14.4	9.3	8.6	6.6
		1.7	2.6	2.6	2.6	1.9
	G15	37.1	23.0	16.9	16.0	12.9
		1.1	3.4	4.8	[illegible]	3.8
B	G5	16.3	3.3	2.4	[illegible]	2.8
		1.5	0.8	0.7	[illegible]	0.8
	G10	30.0	13.0	8.3	7.7	5.8
		1.4	1.5	1.7	1.7	0.9
	G15	36.1	21.6	15.0	14.3	11.8
		1.6	1.9	3.0	3.3	2.8
C	G5	18.8	6.7	4.8	4.8	6.4
		2.2	2.2	2.1	2.3	1.7
	G10	31.4	18.1	14.0	13.9	13.6
		2.6	4.2	4.5	4.6	3.9
	G15	37.5	26.0	21.7	21.6	19.8
		2.0	3.7	3.9	3.8	4.5
D	G5	16.3	6.2	4.7	4.7	6.7
		1.4	1.4	1.4	1.4	1.2
	G10	30.8	17.8	14.5	13.8	15.5
		2.6	4.2	4.5	4.6	3.9
	G15	36.9	25.1	20.9	20.2	19.5
		2.3	4.7	5.2	5.4	4.8

Table 3. Percentage incorrect classification based on 10 trials. See text for explanation of rules. Error U8 refers to a uniform error distribution taking values from -8 to +8. Error G5 refers to a Gaussian error distribution with expected value 0 and standard deviation 5. The other error entries are defined similarly. Data set A is the image indicated by 1's in Fig. 6(a) and B is its complement. Data C is the image of Fig. 6(b), and D is its complement.

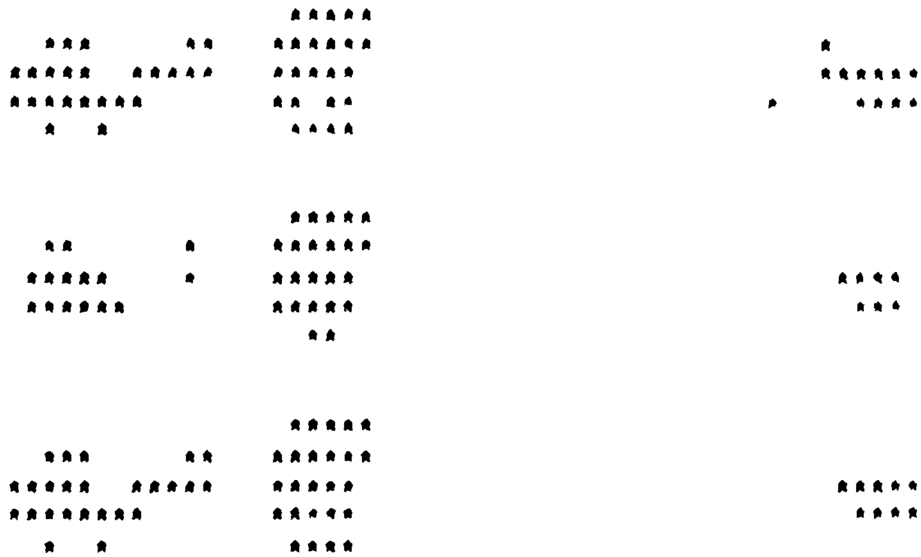


Fig. 5. Top portion is original image. Next portion is result of 5/9 rule, and bottom is result of modified 5/9 rule.

```
000000000000000000000000000000
000000000000000000000000000000
000000000000000000000000000001
000000000000000000000000000011
000000000000000000000000000011
000000000000000000000000001111
000000000000000000000000011111
000000000000000000000000011111
000000000000000000000000111111
000000000000000000000111111111
000000000000000000011111111111
000000000000000011111111111111
000000000000000011111111111111
000000000000001111111111111111
000000000000001111111111111111
000000000000011111111111111111
000000000011111111111111111111
000000000011111111111111111111
000000000111111111111111111111
000000000111111111111111111111
000000000111111111111111111111
000000000111111111111111111111
000000000111111111111111111111
000000001111111111111111111111
000000001111111111111111111111
000000001111111111111111111111
000000000111111111111111111111
000000000111111111111111111111
0000000C)011111111111111111111
000000000011111111111111111111
```

Fig. 6(a)

```
0000000000000000000000000
0000000000000000000000000
0000000000000000000000000
0000000000000010000000000
0000000000001111000000000
0000000000001111000000000
0000000000011111000000000
0000000011111111000000000
0000001111111111100000000
0000001111111111110000000
0000111111111111110000000
0000111111111111110000000
0000111111111111111000000
0001111111111111111100000
0000111111111111111100000
0000111111111111111100000
0000111111111111111000000
0000001111000011111000000
0000000000000000000000000
0000000000000000000000000
0000000000000000000000000
0000000000000000000000000
0000000000000000000000000
```

(b)

5. Some monotone equivariant segmentation techniques. Many commonly used spatial techniques for image enhancement lead to monotone equivariant segmentation algorithms. To illustrate this, consider the histogram modification techniques described in [3], pp. 127-136. If the data are first rank ordered, and after application of the technique, the output is labeled by the original data values, the resulting technique is easily seen to be monotone equivariant. For example, if the original data values are 2.3, 5, 8.7, 12 and 20, the rank-ordering would produce ranks of 1 through 5. If a histogram modification technique produces output levels of 3 and 5, and if this ouput is then labeled as 8.7 and 20 (the values of the original data corresponding to these ranks), then the resulting technique is monotone equivariant.

We described in [9] a segmentation technique that is also monotone equivariant. It will be convenient to briefly describe it here. The input data is an m by n array A of nonnegative integers. The program consists of the following parts:

A. Prefiltering. A version of a k by k mean or median filter is applied to smooth the data, and the ouput is rounded to the nearest integer to form a matrix B.

B. Thinout. A frequency count is made of the values appearing in B. Those values that occur with frequency less than some threshhold are deleted and the points corresponding to them reassigned to the closest valid remaining data value. Alternately, the k highest occurring values can be retained. The resulting matrix is denoted C.

C. Suppose matrix C has data values $j_1, j_2, \ldots, j_k$. The 3 by 3 <u>dispersion</u> of value j_1 is defined as follows: For each point in C having value j_1, look at a 3 by 3 neighborhood centered on that point, and calculate the number of points in that neighborhood that do not have value j_1. Compute the average over all points having value j_1 for which 3 by 3 neighborhoods can be found. This is the dispersion for j_1. Continue the process for $j_2, j_3, \ldots, j_k$.

D. Various options exist for deleting 1 or more of the data values having the highest level dispersion. The simplest is to just delete the highest level value. Other options might include deleting all values whose dispersion exceeds 0.9 on the first pass, and then dropping this cutoff down by increments of 0.1 until a stopping criterion is reached. Various options also exist for the reassignment of points whose value has been deleted. They proceed on both a global and local basis. Cluster means can be computed for each valid cluster, and points assigned either to the nearest mean, or by means of a Bayes decision rule distance. Alternately, this can be done on a global basis, and all points having a given deleted value can be assigned to the nearest valid cluster. These techniques are called ANEAR and ANEARU. A second type of technique involves assignment of points according to whether their labels are above or below the averages of the next valid clusters that are above or below them. This can be done on either a local or global basis, and the resulting techniques are ANEXT and ANEXTU. These techniques as they have been described are not monotone equivariant. If, however, the input data are first rank ordered, and the output is labelled according to which input value corresponds to which rank, then the resulting techniques do become monotone equivariant. An illustration of the operation of the algorithm occurs in Table 4. The input data consists of a mixture of 2 normal distributions, one with expected value 30 and the other with expected value 20, each having a standard deviation of 4. The data entries are rounded to the nearest integer, and a 3 by 3 mean prefilter is used. This prefilter tries to remove outliers by deleting the highest and lowest entries from each neighborhood, and outputting the average of the remaining 7 members. The "nondirected dispersion" shown in this table is identical to the version of dispersion we described earlier. The "directed dispersion" refers to the average of the maximum number of points in the subregions to the North, South, East and West of the center point that lie in regions other than that of the center point. The idea here is that the directed dispersion would distinguish between boundary points to a region and noise. Notice the dramatic drop in the values of the dispersion when the proper number of regions is found. The segmentation algorithm was ANEAR, and the input data was arranged so that one population was in the upper half of the input matrix, and the other in the lower half.

MIXTURE OF 2 NORMAL DISTRIBUTIONS WITH EXPECTED VALUES OF 20 AND 30. OUTPUT OF NON-DIRECTED AND DIRECTED DISPERSIONS AT VARIOUS STAGES OF THE CLUSTERING PROCESS

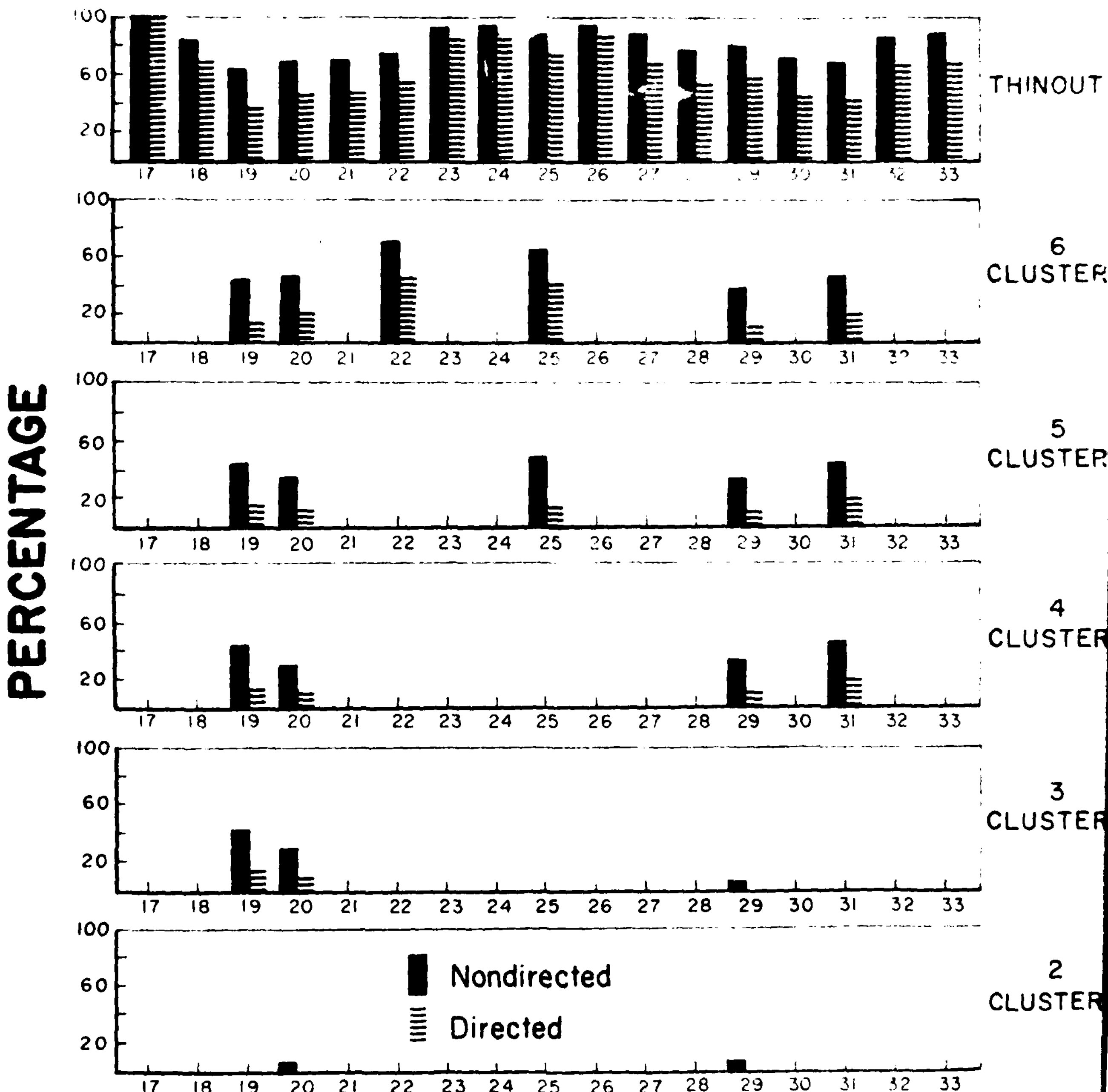


Table 4.

7. Construction of a segmentation algorithm. A program based on the earlier material will now be described and illustrated. The input data is as in Section 3. The program itself has two phases:

Phase 1. Find levels at which to slice the data.

Phase 2. Make appropriate slices and use a j/k rule to increase the probability of correct classification.

As was noted in section 4, the Phase 2 portion of the algorithm is a flat segmentation technique. The program can be used as a cleaning algorith, or it is suited for the determination of the principal regions of a digital picture.

Description of Phase 1. 6 methods of determining the slice levels are built into the program.

Method 1. The mean and standard deviation of the input data are determined. The slices are taken at M-2S, M-S, M, M+S, and M+2S, where M is the mean and S the standard deviation. If this get's one outside the data range, then the extreme values are modified so that they lie between the minimum and maximum values of the input.

Method. 2. A frequency count is taken of the levels that appear in the input data, and slices are taken at relative maximum points of this distribution.

Method. 3. This uses relative minima of the dispersion.

Method 4. This is the same as method 3, except that if A represents the vector of distinct data values, then for A[i], one is interested in the average number of points that do not lie in A[i-1], A[i) or A[i+1]. One then takes the relative minimum points.

Method 5. This too is similar to Method 3, except that the output is the average variance in t by t neighborhoods centered on points having a specified value. The slice levels again are the relative minima.

Method 6. User specifies slice levels.

B. Description of Phase 2. Having arrived at a list of slice levels. one now chooses a j/k^2 rule. The actual slices are constructed so that they are halfway between the slice input levels, and the resulting slices are given the labels of the slice inputs. Thus if one wanted slices at 17, 20, 22, 26, and 30, one would actually look at the increasing sequence of sets $M_1 \subset M_2 \subset M_3 \subset M_4 \subset M_5$, where $M_i = \{x: \le k_i\}$ with $k_1 = 18.5$, $k_2 = 21$, $k_3 = 24$, $k_4 = 28$, and k_5 = the maximum value of X. This forms 5 clusters, and they are assigned the values 17, 20, 22, 26 and 30. Alternately, they could be labeled by their means.

The actual implementation of this process is very simple. For input data X, one looks at the lowest actual slice level k_1. A Boolean matrix is formed by letting 1 denote points whose value is $\le k_i$ and 0 elsewhere. If a j/k^2 rule is used, one then looks at k by k regions. If the number of 1's in such a region is $\ge j$, the output is 1; otherwise, it is 0. The output for k_2 is added to this, and the process continues until the supply of actual slice levels is exhausted. Computation time may be reduced by observing that if at a given stage a 1 is produced, then all succeeding stages must also produce a 1, so there is no reason to consider neighborhoods involving that point. This of course is simply a reflection of the fact that a j/k^2 rule represents an isotone mapping on $P(X)$. The actual implementation can be further optimized by first applying a k by k median filter, and applying the Phase 1 techniques to the output of this filter.

We close by considering some examples involving real data. The data are from two sources: the Westinghouse FLIR tape and an infrared weather tape that was supplied by the Air Force Geophysical Laboratory at Hanscom AFB. Plates 1-5 illustrate the 5 methods of selecting slice levels that were explained earlier in this section, and Plate 6 shows the output of the SEGMENT program on the same data set. Though there are some minor differences shown between the various j/k^2 rules used, all of the methods did a good job of showing the central object of interest - a tank. One should bear in mind that the shades of grey in the pictures serve only to distinguish regions. They do not necessarily represent actual temperature levels of the regions. It is also of interest to note the difference in resolution between the 13/25 rule and the other rules that are used. This is especially apparent in Plates 14 and 17.

REFERENCES

[1] F. Baluicu, Order theoretic classification of cluster methods, University of Massachusetts doctoral dissertation, 1981.

[2] G. Birkhoff, Lattice theory (3rd edition) Amer. Math. Soc. Colloq. Publ., vol. 25, Providence, 1967.

[3] T. S. Blyth and M. F. Janowitz, Residuation theory, Pergamon Press, Oxford, 1972.

[4] R. C. Gonzales and P. Wintz, Digital image processing, Addison-Wesley, Reading, 1977.

[5] M. F. Janowitz, An order theoretic model for cluster analysis, SIAM J. Appl. Math. 34 (1978), 55-72.

[6] ____________, Semiflat L-cluster methods, Discrete Mathematics 21 (1978), 47-60.

[7] ____________, Monotone equivariant cluster methods, SIAM J. Applied Math. 37 (1979), 148-165.

[8] ____________, Preservation of global order equivalence, J. Math. Psych. 20 (1979), 78-88.

[9] ____________, Cluster analysis package for image segmentation, (to appear).

[10] ____________, Cluster analysis algorithms for image segmentation, University of Massachusetts Technical Report J8102, 1981.

[11] N. Jardine and R. Sibson, Mathematical taxonomy, Wiley, New York, 1971.

[12] D. W. Matula, Graph theoretic techniques for cluster analysis algorithms, in "Classification and Clustering" (J. van Ryzin, ed.) Academic Press, New York, 1977.

Plate 1. Portion of Record 21 of Westinghouse FLIR Tape. Slice program output using Method 1. Upper pictures represent output using iterated 3 by 3 mean filter and 3/5 rule. Bottom pictures represent 5/9 rule and 13/25 rule.

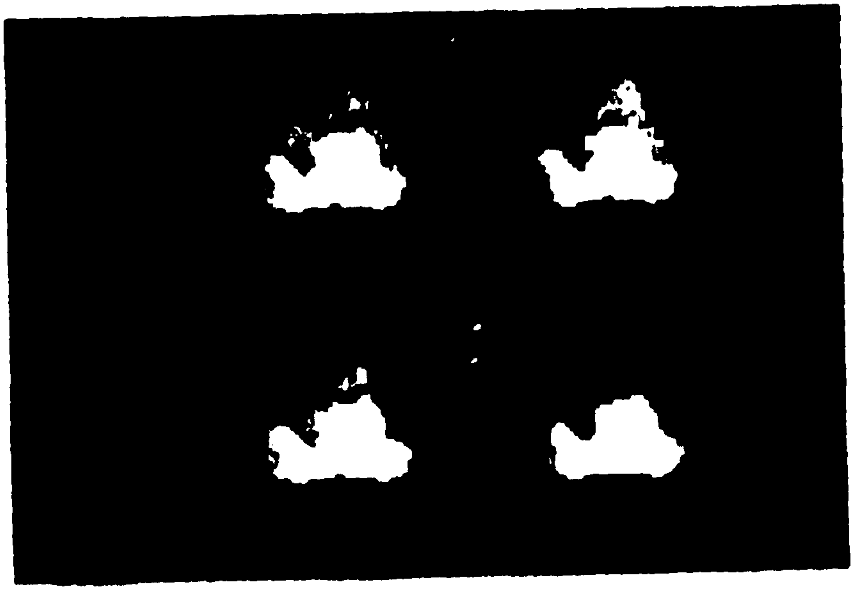


Plate 2. Same as Plate 1, except that slice levels are obtained by Method 2.

Plate 3. Same as Plate 1, except that ... ... are obtained by Method 3.

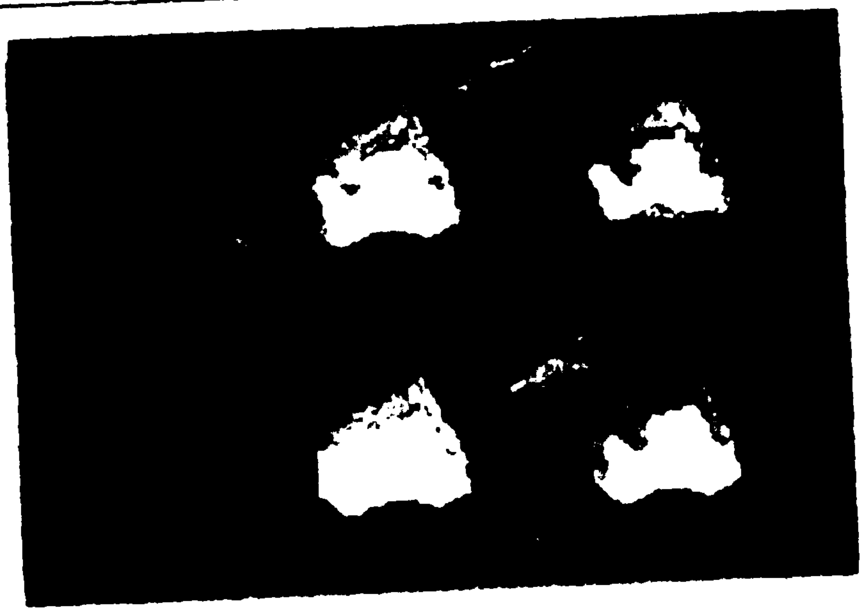


Plate 4. Same as Plate 1, except that ... levels ... obtained by Method 4.

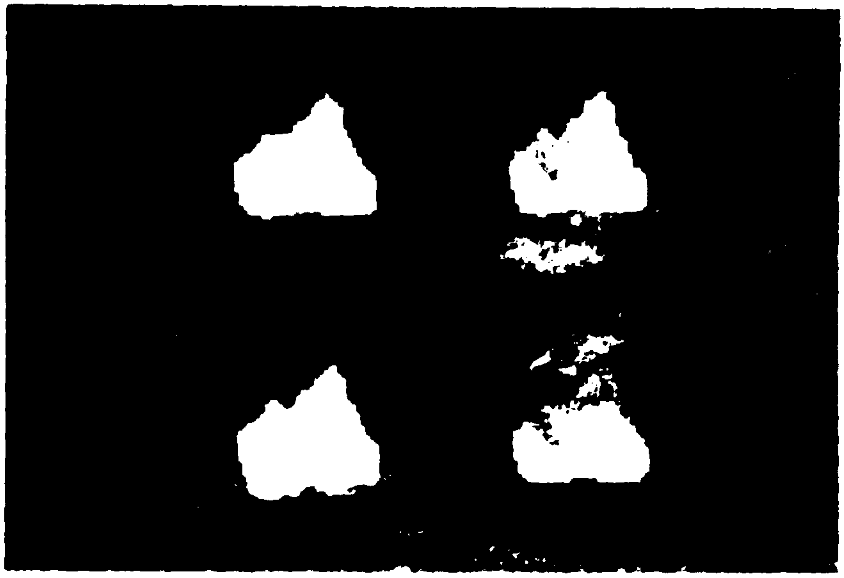


Plate 5. Same as Plate 1, except that slice levels are obtained by Method 5.

Plate 6. Output of Segment program [illegible] Record 21 of FLIR tape. Top row represents original data, output of [illegible] by 3 mean filter, output of Thinout program. Bottom row [illegible] output at 5, [illegible] output levels. The segmentation [illegible] connected 3 by 3 dimension [illegible].

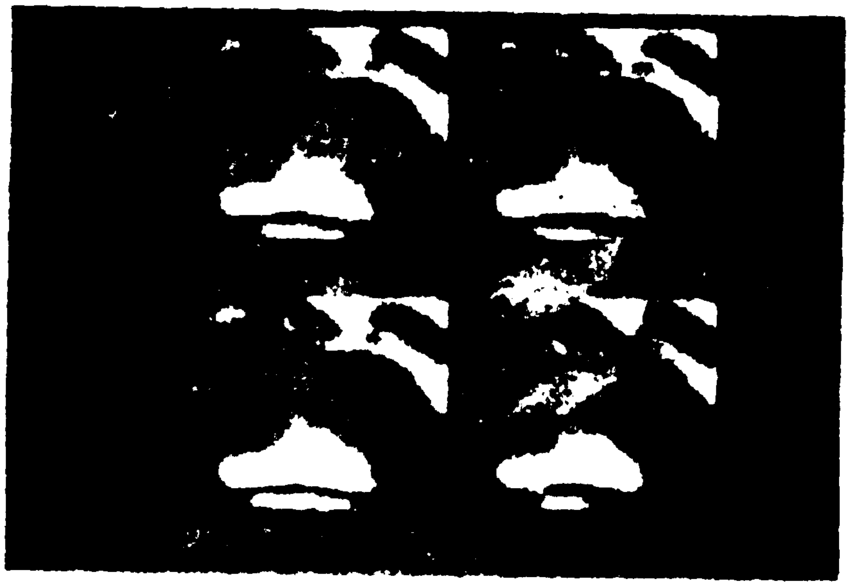


Plate 7. Portion of [illegible] Westinghouse [illegible] tape. Slice program output using Method [illegible]. [illegible] represent output [illegible] iterated 3 by 3 mean filter, [illegible] and 15/25 rule.

Plate 8. Data [illegible] explanation of Method [illegible]

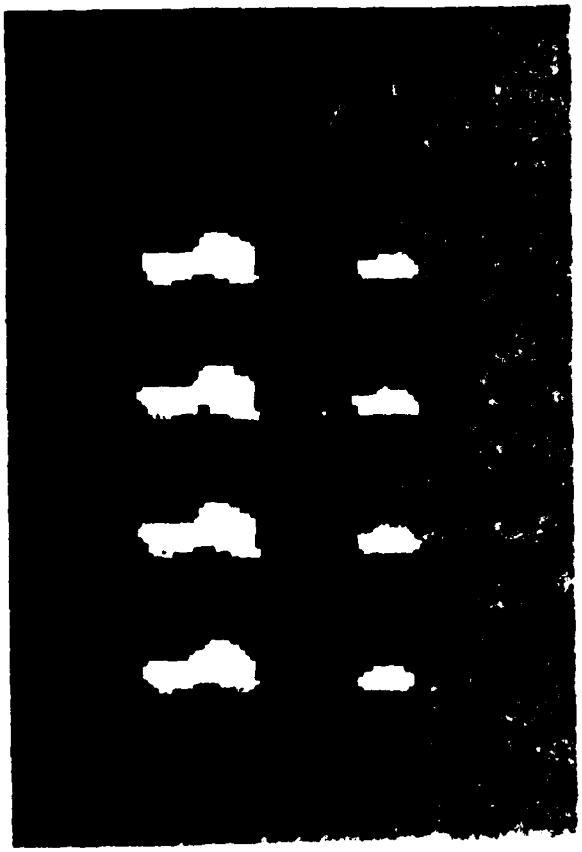
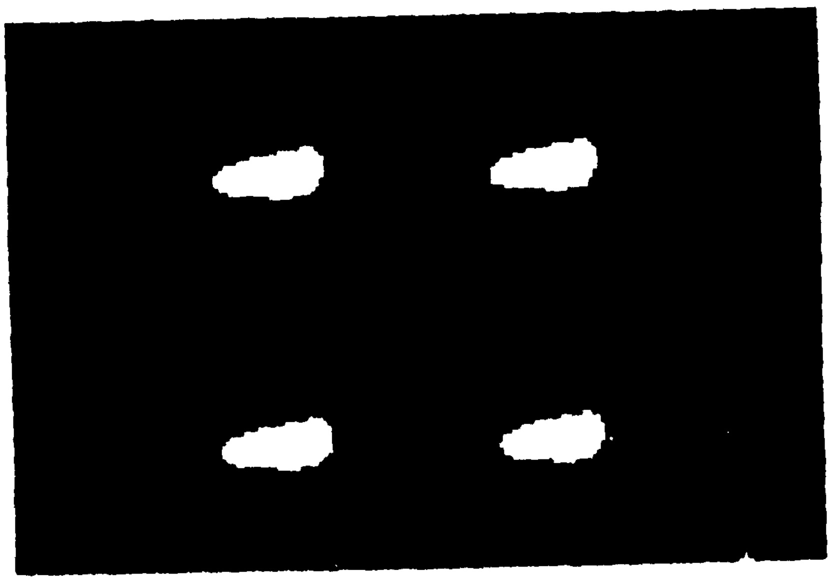


Plate 9. Portion of Record [illegible] tape. Slice pro[illegible] using Method 2. [illegible] under Plate 1 for [illegible] explanation of [illegible]

Plate 10. Same as Plate 9, except that Method 5 is used.

<u>Plate</u> 11. Same as Plate 6, except that data is from Record 68.



<u>Plate</u> 12. Portion of Record 12 of FLIR tape. Slice program output using Method 3. Format follows that of Plate 1.

Plate. 13. Same as Plate 6, except that data is from Record 12.

Plate 14. Portion of weather picture showing Gulf Stream and some cloud cover. Slice program using Method 5. Format follows that of Plate 1.

Plate [illegible]

Plate [illegible] ... area and ...