itself to a fast, highly parallel implementation. Some experiments have shown that this relaxation method performs well under considerable distortion of the measurements but that its efficacy depends strongly on the form of the local compatibility function used. For the type of local compatibility used here it appears that best results are achieved by the use of *maximum* and *minimum* for fuzzy logical OR and AND respectively.

ACKNOWLEDGMENT

The help of Kathryn Riley in preparing this paper is gratefully acknowledged. The author would also like to thank Dr. Azriel Rosenfeld for many stimulating discussions, and the referees for their helpful criticisms.

REFERENCES

[1] H. G. Barrow and J. M. Tenenbaum, "MSYS: A system for reasoning about scenes," Stanford Research Institute, Menlo Park, CA, Tech. Note 121, Apr. 1976.
[2] L. Kitchen and A. Rosenfeld, "Discrete relaxation for matching relational structures," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 12, pp. 869-874, Dec. 1979.
[3] S. Peleg, "A new probabilistic relaxation scheme," Tech. Rep. 711, Computer Science Center, Univ. Maryland, College Park, Tech. Rep. Nov. 1978.
[4] S. Ranade and A. Rosenfeld, "Point pattern matching by relaxation," *Pattern Recognition*, to appear.
[5] W. H. Tsai and K. S. Fu, "Error-correcting isomorphisms of attributed relational graphs for pattern classification," School of Electrical Engineering, Purdue Univ., West Lafayette, IN, Tech. Rep. TR-EE 79-3, Jan. 1979.
[6] L. A. Zadeh, "Fuzzy sets," *Inform. and Control*, vol. 8, pp. 338-353, June 1965.

Asymmetric Clusters of Internal Migration Regions of France

JOHN P. BOYD

Abstract—**The concept of hierarchical clustering is generalized to allow asymmetric input and output. The new method is called "topological clustering" and is computed by the author's program TOPCLU. The result of ordinary clustering can be thought of as a chain of reflexive, symmetric, and transitive relations (i.e., equivalence relations or partitions) ordered by inclusion. Similarly, topological clustering results in a chain of reflexive, transitive, but not necessarily symmetric relations (i.e., preorders or topologies), also ordered by inclusion. This technique is applied to internal migration regions of France where it is particularly natural to look for asymmetric relationships.**

I. INTRODUCTION

Many kinds of data are in the form of square, but asymmetric, matrices. This type of data can be interpreted as the flow of objects, people, or ideas. In psychology, Rosch [6] has shown that similarity judgments are asymmetric in an important way. Thus the usual measures of similarity, such as correlation coefficients, that force symmetry onto the data may unfortunately obscure this important aspect of the phenomenon.

Since hierarchical clustering procedures are designed for the symmetric situation, it is important to find a more general representation of these square matrices which will accurately display asymmetric situations. This correspondence presents a topological clustering approach to this problem. This procedure gives a nested sequence of partitions (as do symmetric clustering routines) but, in addition, gives a partial ordering on the equivalence classes (cluster) of these partitions. The result is for each cutoff level a visual decomposition of the data into its symmetric and asymmetric parts.

This method is applied to Slater's [7] data on the internal migration patterns of France. The result is a clear picture of not only the important clusters but also the pattern of migration between these clusters.

Manuscript received June 28, 1979; revised November 5, 1979.
The author is with the School of Social Sciences, University of California, Irvine, CA 92717.

II. TOPOLOGICAL CLUSTERING

A brief description of the topological clustering method will be given here. The approach in this correspondence is to generalize the single-link (or connectedness) method, following Jardine and Sibson's [4] strategy of modifying the dissimilarity matrix in order to obtain a new matrix satisfying certain additional axioms. This allows a nested sequence of discrete structures to be read off from the new matrix D^*.

It is assumed that one is given a square matrix D whose entries may be interpreted as a set of dissimilarity (or similarity) coefficients. Since similarity coefficients can be easily transformed to dissimilarities by a suitable order reversing function, it will be assumed that D is a dissimilarity matrix of nonnegative real numbers satisfying the single axiom

$$\text{(Definiteness) } D[x, y] = D[y, x] = 0, \qquad \text{iff } x = y$$

for all x, y. Note that neither symmetry nor the triangle inequality is assumed to hold on the data matrix D. This matrix D is to be modified in some minimal way so as to obtain a matrix D^* satisfying the ultrametric inequality described below.

To see how topological clustering generalizes the single-link method, the latter must be described in a certain format. Namely, the single-link method results in a function h from the nonnegative reals into the set of all equivalence relations on the row (=column) labels for the matrix D. This function is required to be increasing with respect to set inclusion of the ordered pairs in each equivalence relation. That is, if $c \leq c'$ then $h(c) \subset h(c')$.

Since an equivalence relation is reflexive, symmetric, and transitive, the generalization to another data structure which can represent asymmetric matrices is both natural and obvious. That is, simply omit the symmetry axiom! Fortunately, reflexive and transitive relations are well-known by the name of *preorders*. Furthermore, just as the set of all equivalence relations on a given set is in a natural one-to-one correspondence with all partitions on the same set, so are preorders in a natural one-to-one correspondence with topologies, at least for finite sets [3]. Therefore this method will be called *topological clustering*. The construction of the topology T_P corresponding to a preorder P on a set X is easily described: $U \subset X$ is open in T_P if and only if for all x in X, and u in U, xPu implies $x \in U$. Conversely, given a topology T on X, the corresponding preorder P_T is defined for all x, y in X by the condition xP_Ty if and only if for all open sets U in T, $y \in U$ implies $x \in U$. These two mappings are inverses of each other if X is finite. Thus preorders and topologies for finite sets are essentially identical. However, since finite preorders are easier to represent visually than their corresponding topologies, this correspondence will opt for the preorder form. Nevertheless, the reader should keep in mind that a preorder is really a topology.

The goal then of topological clustering is to find a function h from the nonnegative reals into the set of all preorders on the row labels. This function, as in the case of symmetric clustering, must

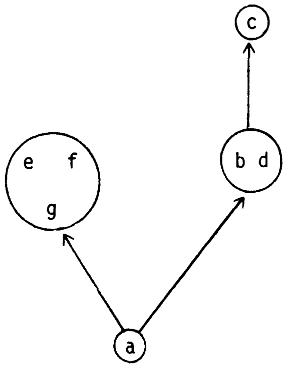


Fig. 1. Diagram of the preorder given by ordered pairs.

$aPa, b, c, d, e, f,$ and g
$bPb, c,$ and d
cPc
$dPb, c,$ and d
$ePe, f,$ and g
$fPe, f,$ and g
$gPe, f,$ and g

Note, for example, that $bE_P d$ holds, but that $bE_P c$ does not. Also, $[a]_P <_P [c]_P$ holds, but $[a]_P \prec [c]_P$ does not. The topology T_P consists of the sets, ϕ, X, $\{a\}$, $\{a, b, d\}$, $\{a, b, c, d\}$, $\{a, e, f, g\}$, and $\{a, b, d, e, f, g\}$.

satisfy the condition that if $c \leq c'$, then $h(c) \subset h(c')$. That is, everything is the same except that the concept "preorder" replaces "equivalence relation." Of course, this "goal" is not very desirable unless it can be shown that any finite preorder can be presented in an attractive and useful format, preferably in such a way as to see what would have happened if the old assumption of symmetry were to be imposed. Fortunately, such a format can be found.

First, note that every preorder P can be decomposed into a symmetric and an antisymmetric part as follows. Define a relation E_P by the rule

$$xE_P y, \qquad \text{iff } xPy \text{ and } yPx.$$

Obviously E_P is an equivalence relation. Let $[x]_P$ denote the E_P-equivalence class containing x. The relation E_P represents the summetric part of P. Next, a partial order $\leq_P$ defined as a reflexive, antisymmetric, and transitive relation, can be defined on the equivalence classes of E_P by the rule

$$[x]_P \leq_P [y]_P, \qquad \text{iff } xPy.$$

This decomposition of preorders can be used to display them effectively. Equivalence relations can be represented by the Venn diagrams of their partitions by drawing circles around the equivalence classes. A partial order $\leq$ is usually shown by first defining the "covering" relation, denoted by $\prec$ by omitting redundant ordered pairs. That is, $x \prec y$, read x is *covered* by y, if and only if

1) $x \leq y$,
2) $x \neq y$, and
3) for no b distinct from both x and y is it true that $x \leq b \leq y$.

The original partial order can be recovered if the underlying set is finite by taking the transitive closure of the covering relation and then adjoining the identity relation. The *diagram* of a partial order is formed by representing each covering pair $x \prec y$ by an arrow from x to y. Thus $x \leq y$ if and only if on the diagram there is a directed path from x to y.

Combining these two notions a preorder P will be pictured as the diagram of $\leq_P$ drawn on the circles representing E_P. For example, Fig. 1 presents the diagram of a small hypothetical preorder on the set $X = \{a, b, c, d, e, f, g\}$. Note that P consists of exactly 23 ordered pairs, but the covering relation has only three ordered pairs on the four equivalence classes. Therefore, significant conceptual and visual economy is effected by this representation, indicating that a preorder may be in fact a desirable, or at least an understandable, goal.

One way to achieve this goal is to lower some of the numbers in the dissimilarity matrix D to form a new matrix D^* such that for every cutoff $c \geq 0$ the relation P_c, defined by

$$xP_c y, \qquad \text{iff } D^*[x, y] \leq c$$

is a preorder. It can be shown that, just as in the symmetric case, P_c is transitive for all $c \geq 0$ if and only if D^* satisfies the

$$\text{(ultrametric inequality) } D^*[x, z] \leq \max\{D^*[x, y], D^*[y, z]\}.$$

An efficient procedure for calculating D^* is given by the three loops:

```
for y := 1 step 1 until n do
for x := 1 step 1 until n do
for z := 1 step 1 until n do
begin
    max := (if D[x, y] ≥ D[y, z] then D[x, y], else D[y, z]);
    if D[x, z] > max then D[x, z] := max
end
```

Note the funny order of the loops. The fact that the middle element y is done first insures that the ultrametric is satisfied after only one pass. The method gives a "subdominant" solution D^* in the sense of Jardine and Sibson [4]. That is, if D' is any other matrix satisfying the ultrametric and the requirement that $D'[x, y] \leq D[x, y]$ for all x, y, then $D'[x, y] \leq D^*[x, y]$ for all x, y. In other words a subdominant solution D^* is the minimal "contraction" of the data matrix that satisfies the ultrametric. Note that the map $D \to D^*$ is a continuous function $(\mathbb{R}^+)^n \times (\mathbb{R}^+)^n \to (\mathbb{R}^+)^n$, a virtue shared by the single-link method but missing from the diameter method [4].

The author's program called TOPCLU which effects this analysis was written in SIMULA, an ALGOL based simulation language that is good for handling sets among other things. It has worked fine for data from a 60 by 60 matrix and can handle in its present format a 99 by 99 matrix. The program takes as input a square matrix D of real numbers, converts them to dissimilarities if necessary, computes the matrix D^* satisfying the ultrametric, and prints out D^*. Then it asks the user what cutoffs c are desired and prints out P_c as described above.

It is interesting to compare TOPCLU with a program, DECOMP, written by Brams [1]. DECOMP takes as input a binary matrix and forms the transitive closure, printing out the resulting preorder, plus some row and column statistics. Except for these statistics, the author's program would do the very same thing on a matrix of zeros and ones if the cutoffs were between zero and one. Therefore, the author's program generalizes Bram's procedure, which only takes 0–1 data, to a procedure accepting any matrix of nonnegative real numbers.

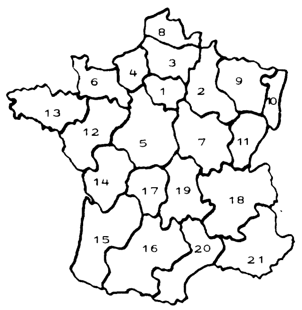
TOPCLU itself does not normalize matrices. The author recommends that only normalized matrices be used as input, unless there are other grounds for believing that the asymmetries are not merely caused by a row and column size bias.

III. French Migration

Slater [7] presented data on the internal migration patterns among 21 regions of France (see Fig. 2). He normalized the matrix so that each row and each column added to 1000, using an itera-

TABLE I
ADJUSTED 1962–1968 FRENCH INTERREGIONAL MIGRATION

	Destination																				
Origin	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0	49	86	76	66	78	40	44	23	16	23	56	81	46	54	47	53	29	29	42	61
2	42	0	132	23	31	23	114	55	179	46	101	25	28	29	26	21	17	29	15	26	38
3	60	178	0	125	31	36	22	302	29	17	14	16	18	16	19	23	16	22	19	16	22
4	58	17	158	0	60	235	14	61	21	17	18	49	77	32	23	16	41	22	21	24	37
5	72	25	33	65	0	63	77	37	26	25	22	100	51	115	41	35	78	22	67	17	28
6	82	16	37	212	65	0	20	25	20	26	10	117	182	33	19	14	47	17	21	21	18
7	51	91	32	31	60	16	0	28	35	31	211	19	23	23	25	19	16	131	87	34	38
8	51	85	231	80	33	41	36	0	49	34	34	29	24	27	40	36	24	39	31	36	43
9	25	153	28	25	23	13	44	64	0	221	136	15	25	22	28	28	14	33	21	30	54
10	27	60	17	26	38	16	39	59	235	0	144	19	39	22	34	35	6	39	25	57	64
11	27	91	10	24	17	20	214	22	114	189	0	9	20	24	23	18	7	90	14	19	50
12	53	22	18	41	96	154	16	20	19	16	14	0	206	163	25	25	39	17	16	19	20
13	90	28	32	94	58	139	15	38	33	18	19	238	0	39	37	14	11	22	15	18	40
14	49	14	23	32	117	19	15	27	30	44	22	129	45	0	145	36	149	20	31	22	30
15	48	18	23	28	40	27	15	30	30	26	22	37	27	145	0	190	126	29	38	57	44
16	47	19	29	21	31	18	25	26	25	71	13	22	18	33	188	0	89	37	47	185	56
17	55	24	21	16	81	20	15	27	13	8	16	27	18	144	126	69	0	24	234	40	21
18	31	28	22	17	28	24	129	28	31	56	86	19	13	11	24	47	34	0	151	96	125
19	41	20	23	16	69	27	92	23	22	26	25	30	15	22	35	70	198	144	0	53	49
20	44	25	25	19	27	8	27	42	32	50	32	18	23	20	46	202	15	110	74	0	162
21	47	36	20	29	29	23	31	43	34	65	38	25	67	33	43	57	20	126	45	190	0



1 Région Parisienne	12 Pays de La Loire
2 Champagne	13 Bretagne
3 Picardie	14 Poitou-Charentes
4 Haute-Normandie	15 Aquitaine
5 Centre	16 Midi-Pyrénées
6 Basse-Normandie	17 Limousin
7 Bourgogne	18 Rhône-Alpes
8 Nord	19 Auvergne
9 Lorraine	20 Languedoc
10 Alsace	21 Provence-Côte d'Azur-Corse
11 Franche-Comté	

Fig. 2. Twenty-one French regions.

tive procedure recommended by Romney [5] and others [2] (see Table I). The advantage of normalization for asymmetric clustering is that it removes the size of the row or column total as a false asymmetric bias. That is, making the rows, but not the columns, add up to 1000 does not remove the column bias that one would expect from the "null hypothesis" that the entries within each column be identical. Thus the conditional probability of migration is not a good way to detect an asymmetric bias.[1] For example, the fact that a higher percentage of people in Champagne move to Paris than vice versa does not necessarily indicate a greater inclination on the part of the average resident of Champagne to move to Paris. It may only indicate that Paris is more populous than Champagne. Therefore, to remove the size bias, one forces the columns as well as the rows to add up to 1000, resulting in a doubly stochastic matrix (if you divide by 1000). This is done by alternately normalizing the rows and then the columns until the matrix converges to a doubly stochastic matrix D. This process can be shown to converge if there are no zero rows or columns. Furthermore, the cross-product ratio is preserved:

$$D_{ij} D_{i'j'} - D_{i'j} D_{ij'} = M_{ii} M_{jj} - M_{ij} M_{ji}$$

for all i, i', j, j'. Returning to the Champagne–Paris example, one can see, in fact, from Table I that after the size bias is removed by the normalization, there is a slight tendency for people to be more inclined to move from Paris to Champagne (49 versus 42 per 1000 emigrants). These numbers were transformed to dissimilarities D by subtracting them from 302, the maximal entry. Slater did a symmetric hierarchical clustering of these regions and then discussed separately some of the asymmetric linkages. Topological clustering enables one to do both simultaneously. Table II gives the program's output matrix D^*, satisfying the ultrametric inequality. The preorders for three cutoffs are shown in Fig. 3. Note that the Venn diagrams for the equivalence classes are not really circles, but outlines of the geographical regions. The arrows of the covering relation are oriented to correspond to the direction of net migration.

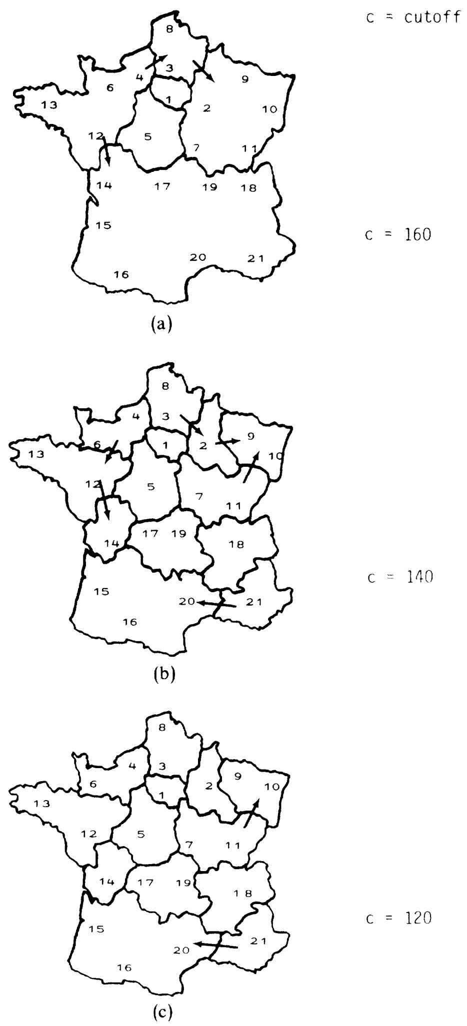
A word of warning should be inserted here. Since this method is a generalization of the single-link method, it shares its most controversial property, "chaining." That is, if there is a chain x, y, z, where x is close to y and where y is close to z, then it will put x close to z, even though they may be quite distant in fact. Sometimes long chains are formed as cluster when the ends are similar. These clusters may fail to be compact. Although it can be

[1] Similar remarks, of course, could be made for other kinds of data. For example, word associations tend to go to the most common words. Normalization would remove this bias.

TABLE II
ULTRAMETRIC FOR DATA OF TABLE I AFTER THEY HAVE BEEN TRANSFORMED 1) INTO DISSIMILARITY BY $d = 302 - s$ AND THEN 2) BY THE ULTRAMETRIC ALGORITHM

THE ULTRAMETRIC

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216	216
2	212	0	170	173	185	173	158	170	123	123	158	173	173	171	171	171	171	171	171	171	171
3	212	124	0	173	185	173	158	0	124	124	158	173	173	171	171	171	171	171	171	171	171
4	212	144	144	0	185	67	158	144	144	144	158	120	120	139	157	157	153	158	153	157	157
5	212	187	187	187	0	187	187	187	187	187	187	187	187	187	187	187	187	187	187	187	187
6	212	144	144	90	185	0	158	144	144	144	158	120	120	139	157	157	153	158	153	157	157
7	212	149	170	173	185	173	0	170	113	113	91	173	173	171	171	171	171	171	171	171	171
8	212	124	71	173	185	173	158	0	124	124	158	173	173	171	171	171	171	171	171	171	171
9	212	149	170	173	185	173	158	170	0	81	158	173	173	171	171	171	171	171	171	171	171
10	212	149	170	173	185	173	158	170	67	0	158	173	173	171	171	171	171	171	171	171	171
11	212	149	170	173	185	173	88	170	113	113	0	173	173	171	171	171	171	171	171	171	171
12	212	148	148	148	185	148	158	148	148	148	158	0	96	139	157	157	153	158	153	157	157
13	212	148	148	148	185	148	158	148	148	148	158	64	0	139	157	157	153	158	153	157	157
14	212	173	173	173	185	173	173	173	173	173	173	173	173	0	157	157	153	158	153	157	157
15	212	173	173	173	185	173	173	173	173	173	173	173	173	157	0	112	157	158	157	117	140
16	212	173	173	173	185	173	173	173	173	173	173	173	173	157	114	0	157	158	157	117	140
17	212	173	173	173	185	173	173	173	173	173	173	173	173	158	158	158	0	158	68	158	158
18	212	173	173	173	185	173	173	173	173	173	173	173	173	158	158	158	151	0	151	158	158
19	212	173	173	173	185	173	173	173	173	173	173	173	173	158	158	158	104	158	0	158	158
20	212	173	173	173	185	173	173	173	173	173	173	173	173	157	114	100	157	158	157	0	140
21	212	173	173	173	185	173	173	173	173	173	173	173	173	157	114	112	157	158	157	112	0



1 Région Parisienne
2 Champagne
3 Picardie
4 Haute-Normandie
5 Centre
6 Basse-Normandie
7 Bourgogne
8 Nord
9 Lorraine
10 Alsace
11 Franche-Comté
12 Pays de la Loire
13 Bretagne
14 Poitou-Charentes
15 Aquitaine
16 Midi-Pyrénées
17 Limousin
18 Rhône-Alpes
19 Auvergne
20 Languedoc
21 Provence-Côte d'Azur-Corse

Fig. 3. Preorders on 21 French regions for three cutoff values.

argued, especially in biology [4], that chaining is desirable, it is regarded as a weakness in most social science circles. In this data, however, chaining does not occur to any noticeable degree. Furthermore, one can easily conceive of a diameter method version of topological clustering so as to avoid this problem.[2]

Nevertheless, the reader should be aware of the chaining property in order to properly interpret the diagrams. For example, in Fig. 3(a). the flow from the cluster containing Haute-Normandie (4) to the cluster containing Picardie (3), together with the migration from Picardie (3) to the Champagne cluster (2), does *not* imply a direct flow from Haute-Normandie (4) to Champagne (2), even though the transitive closure would produce such an arrow. In this sense the "cover" relation is closer to the data than is the partial order used to define it. To be precise, for a given cutoff $c \geq 0$, $[x]_c \prec [y]_c$ if and only if there is an ordered pair (x', y') equivalent to x and y, respectively, such that $D[x', y'] \leq c$.

Another property of the single-link method is that, for the high cutoff values when there are just two clusters remaining, these two clusters are often very different in size. Such is the case here, where with a cutoff of 200 the Région Parisienne is alone in one cluster and the rest of France is in another. The diameter method would have given two clusters of about equal size. It seems, however, that the cultural and political realities in France are well represented by putting Paris in a class by itself.

Another cautionary note peculiar to the asymmetric aspects of this method is that it is possible to pick a cutoff c such that for some x, y

$$D^*[x, y] < c < D^*[y, x]$$

and where $[x]_c \neq [y]_c$, and yet the difference $D^*[x, y] - D^*[y, x]$ is really so small as to be insignificant. In the absence of a statistical theory of error, one has to guess whether these differences are large enough to seem interesting.

Having made these caveats, however, the picture presented in Fig. 3 is reasonable. The clusters themselves agree almost perfectly with Slater's results. The arrows to the Alsace and Lorraine regions reflect the heavy industrial growth in those areas. Conversely, the movement out of the northwest is consistent with the relative economic backwardness of those regions. Finally, the lack of any net movement in or out of the Région Parisienne indicates that this area may have reached a steady state. This result is somewhat surprising to those familiar with the rapid influx into the capital cities of developing countries such as Mexico or even Russia. Moreover, even in the steady state the great cities of antiquity could not maintain their own populations without a steady flow of people from the rural areas with their higher birthrate. In France, however, these factors are negated by the low birthrate in the rural areas, good health facilities in the cities, as well as industrialization outside of the Région Parisienne.

IV. Conclusion

Topological clusters can be of use to researchers or planners studying the interrelationships among a set of regions or nodes. It is recommended that asymmetric data be normalized so as to remove asymmetries due purely to row and column effects. It is believed that the combination of these two techniques can lead to interesting results that can be easily displayed.

Other topological clustering methods can be constructed that generalize at least some of the extant symmetric methods in addition to the single-link version used in this paper. Thus many of the controversies among the various methods will also be transported into the asymmetric domain of topological clustering. On the other hand, if the idea of topological clustering becomes widely accepted, then clustering methods which do not generalize to it may fall into disfavor even in the symmetric case.

[2] Here is the diameter method. Assume the same program as above except for replacing the last "if" statement with the following:

```
if D[x, z] > max then
begin if D[x, y] > D[y, z],
      then D[x, y] := D[x, z],
      else D[y, z] := D[x, z],
end.
```

That is, instead of contracting the $D[x, z]$ as in the single-link method, one stretches the longest leg until it matches $D[x, z]$. Note that this method involves breaking ties arbitrarily and then treating the two legs differently. This is what causes the discontinuities in the diameter method here and in the symmetric case.

References

[1] S. J. Brams, "DECOMP: a computer program for the condensation of a directed graph and the hierarchical ordering of its strong components," *Behavioral Science*, vol. 13, pp. 344–345, 1968.

[2] W. E. Deming and F. F. Stephan, "On a least squares adjustment of a sampled frequency table when the expected marginal totals are known," *Annals Math. Stat.*, vol. 11, pp. 427–444, 1940.

[3] H. Herrlich and G. E. Strecker, *Category Theory*. Boston: Allyn and Bacon, 1973.

[4] N. Jardine and R. Sibson, *Mathematical Taxonomy*. New York: Wiley, 1971.

[5] A. K. Romney, "Measuring endogamy," in *Explorations in Math. Anthropology*, P. Kay, Ed. Cambridge: MIT Press, 1971.

[6] E. Rosch, "Cognitive reference points," *Cognitive Psychology*, vol. 7, pp. 532–547, 1975.

[7] P. B. Slater, "Hierarchical internal migration regions of France," *IEEE Trans. Syst., Man, Cybern.*, vol. 6, pp. 321–324, Apr. 1976.

Some Experiments in Point Pattern Matching

DARYL J. KAHL, AZRIEL ROSENFELD, FELLOW, IEEE, AND ALAN DANKER

***Abstract*—Given two pictures of a scene taken by different sensors or at different times, one way of matching the two pictures is to extract a set of distinctive local features from each, and then match the resulting point patterns. The sensitivity of point pattern matching to various types of noise and distortion, including omissions and additions, random walks, rotation and rescaling, as well as the use of different feature detection operators to extract the points is investigated. The effects of additional information (e.g., feature types) in overcoming the effects of noise is also studied.**

I. Introduction

Matching two pictures of the same scene is a common problem in computer vision and image processing. This problem arises in connection with registering pictures obtained by different sensors, or by the same sensor at different times. In such situations the occurrence of systematic gray level differences between the pictures, as well as geometrical distortions, often makes it impractical to use conventional correlation techniques for matching. A possible alternative is to segment the two pictures into regions and attempt to pair off the corresponding regions based on their properties (color, size, shape, etc.) [1]; but this approach depends on the reliability of the segmentation process.

Manuscript received December 20, 1978; revised March 13, 1979 and November 5, 1979. This work was supported by the Defense Advanced Research Projects Agency and U.S. Army Night Vision Laboratory under Contract DAAG-53-76C-0138 (DARPA Order 3206).

D. J. Kahl was with the Computer Science Center, University of Maryland, College Park, MD 20742, on leave from IBM Glendale Laboratory, Endicott, NY 13760.

A. Rosenfeld and A. Danker are with the Computer Science Center, University of Maryland, College Park, MD 20742.

0018-9472/80/0200-0105$00.75 © 1980 IEEE