AD-779 331

OBJECTIVE METHODS FOR DETERMINING CLOUD MOTIONS FROM SATELLITE DATA

D. J. Hall, et al

Stanford Research Institute

Prepared for:

Environmental Prediction Research Facility

November 1973

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER: EPRF Technical Report 2-74 (SRI)

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER: AD-779 331

4. TITLE (*and Subtitle*): Objective Methods for Determining Cloud Motions from Satellite Data

5. TYPE OF REPORT & PERIOD COVERED

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): D. J. Hall, F. K. Tomlin and D. E. Wolf

8. CONTRACT OR GRANT NUMBER(s): N62306-71-C-0068

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Stanford Research Institute
Menlo Park, CA 94025

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS:
Commander, Naval Air Systems Command
Department of the Navy
Washington, D.C. 20361

12. REPORT DATE: November 1973

13. NUMBER OF PAGES: 45

14. MONITORING AGENCY NAME & ADDRESS (*if different from Controlling Office*):
Environmental Prediction Research Facility
Naval Postgraduate School
Monterey, CA 93940

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

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151

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

Enhancement of cloud motion computation method
Geostationary meteorological satellite data manipulation
Man-machine interactive computer system
Multilayered cloud pattern motions
Automated cloud motion vectors

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

The PDP 11 computer, with a Vector General display interface, has been programmed to allow the generation of information representing artificial clouds using a cluster specification language. This language can also be used as input to a similar program that can be run on the CDC 6400 computer to generate identical data and later track the identical data. A cloud-motion vector field is then applied to the artificial clouds and they

20. (continued)

are assumed to move into a new position at a later time. Thus we are able to produce artificial cloud pictures having some c the main features of real pictures.

In general, our problem is to vary the known and controlla parameters of the artificial clouds and wind-vector fields so that we can test various features of the cloud motion programs we have developed so far. In this way we discover the limits of their performance, which enables us to make improvements to the cloud-motion recognition programs.

The methods developed make use of infrared information about cloud heights. Computer results--when compared with human tracking of the same data--show that the computer programs can track many cases accurately, even when the cloud motions are in opposing directions for differing height layers. Further work is suggested to simulate real satellite cloud pictures more closely, allowing for controlled improvements to the computer programs.

STANFORD RESEARCH INSTITUTE
Menlo Park, California 94025 · U.S.A.

Scientific Report *November 1973*

OBJECTIVE METHODS FOR DETERMINING CLOUD MOTIONS FROM SATELLITE DATA

By: D. J. HALL F. K. TOMLIN D. E. WOLF

Prepared for:

SATELLITE APPLICATIONS DEPARTMENT
ENVIRONMENTAL PREDICTION RESEARCH FACILITY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93940

CONTRACT N62306-71-C-0068

SRI Project 1005

Approved by:

R. T. H. COLLIS, *Director*
Atmospheric Sciences Laboratory

R. L. LEADABRAND, *Executive Director*
Electronic and Radio Sciences Division

ABSTRACT

The PDP 11 computer, with a Vector General display interface, has been programmed to allow the generation of information representing artificial clouds using a cluster specification language. This language can also be used as input to a similar program that can be run on the CDC 6400 computer to generate identical data and later track the identical data. A cloud-motion vector field is then applied to the artificial clouds and they are assumed to move into a new position at a later time. Thus, we are able to produce artificial cloud pictures having some of the main features of real pictures.

In general, our problem is to vary the known and controllable parameters of the artificial clouds and wind-vector fields so that we can test various features of the cloud motion programs we have developed so far. In this way we discover the limits of their performance, which enables us to make improvements to the cloud-motion recognition programs.

The methods developed make use of infrared information about cloud heights. Computer results--when compared with human tracking of the same data--show that the computer programs can track many cases accurately, even when the cloud motions are in opposing directions for differing height layers. Further work is suggested to simulate real satellite cloud pictures more closely, allowing for controlled improvements to the computer programs.

CONTENTS

ABSTRACT iii

LIST OF ILLUSTRATIONS. vii

LIST OF TABLES ix

I INTRODUCTION. 1

II OBJECTIVES. 5

III THEORY OF AUTOMATIC CLOUD TRACKING. 7
  A. Axioms of Cloud Tracking Theory. 7
  B. Definitions of ISODATA and MOTION Programs 9

IV EXPERIMENTAL METHOD 13
  A. Interpretation of Line-Printer Plots 15
  B. Example of Experimental Methods. 17

V EXPERIMENTS USING PDP 11 AND CDC 6400 COMPUTERS 23
  A. The Intertropical Convergence Zone 23
  B. Three Successive Artificial Pictures of the ITCZ 25
  C. Verification of Artificial Motions by SRI Cloud Console. 36

VI CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK. 41

REFERENCES 43

ACKNOWLEDGMENTS. 45

Preceding page blank

ILLUSTRATIONS

1 A Printer Plot of Clouds of Artificial Data 16

2 Printer Plot of the Successive Picture of Artificial Data 19

3 Cloud Motion Vectors Plotted on a Line Printer 20

4 Motion Vectors Derived by Plotting the Generating-Language Data 24

5 39 Clusters of the ITCZ as Generated from the Data of Table 2 27

6 The Second Picture in the Sequence for the ITCZ 28

7 The Third Picture in the Sequence for the ITCZ 30

8 Photograph of the PDP 11 CRT Screen During Display of the ITCZ Data 31

9 Print Plot with ISODATA Cluster Centers Superimposed 33

10 Vectors Resulting from Pairing of Clusters with the Same Cluster Number 34

11 Motion Vectors After Four Iterations of the MOTION Program, Pictures 1 and 2 35

12 Motion Vectors After Four Iterations of the MOTION Program, Pictures 2 and 3 37

13 Human Tracking of the Artificial Clouds on the SRI Cloud Console 39

Preceding page blank

TABLES

1 Computer Commands for the Clustering Language 14

2 Computer Commands for Generating 10 Clusters of Artificial Data 18

3 Computer Commands for Generating 39 Clusters Simulating the ITCZ 26

4 The Data for Generating Picture 2 29

Preceding page blank

I INTRODUCTION

In this introductory section of the report, we present the context of the research, and mention the background and some prior developments that led to the current studies. Two previous reports have been issued under the same contract as this current work: "Objective Methods for Registering Landmarks and Determining Cloud Motions from Satellite Data,"[1] and "Further Development of Objective Methods for Registering Landmarks and Determining Cloud Motions from Satellite Data."[2] In the title of this report, we have omitted reference to "registering landmarks" since it was not part of the proposal objectives.

Other Stanford Research Institute (SRI) projects closely related to this work are described in the reports entitled: "A Pattern Recognition Technique for Determining Cloud Motions from Sequences of Satellite Photographs,"[3] also funded by the Navy; and, "Development of New Pattern Recognition Methods,"[4] funded by the Air Force. Another related project report for NASA is entitled: "Development of an Automatic Computer System for Measuring Cloud Motion Vectors."[5]

All of the work referred to above has been carried out in collaboration by personnel of the Atmospheric Sciences Laboratory and the Information Sciences Laboratory of SRI. This project for the Navy has been monitored by Mr. Roland E. Nagle, head of the Satellite Applications Department of the Environmental Prediction Research Facility (EPRF) at Monterey, California.

Before we state the specific purpose and goals of this project in the following sections, we will show how our project fits into a more global research program having general benefits related to environmental prediction.

Weather Prediction and Satellite Observations

It is evident that the ability to predict the weather accurately has benefits in Naval operations. One of the primary variables in weather prediction is the field of atmospheric motions over the globe. The field of atmospheric motions cannot be measured directly--only by the effect of the wind on an object, such as a balloon. Since it is expensive to send up enough balloons to cover the vast areas of interest in global weather prediction, it is fortunate that nature provides us with "balloons" of visible water vapor, in the form of clouds. These "balloons" are not stable in shape or size; therefore, tracking them over long periods or with a limited number of observations is not always easy.

Technological advances have recently provided us with weather satellites, giving spectacular pictures of the clouds and earth. Our general task is to make use of these pictures to track the cloud motions so that these motions may be used in numerical models for prediction of the weather. We are not implying that cloud motions and wind motions are identical--only that we can observe cloud motions but we cannot observe wind motions in satellite pictures.

Satellite observations of the earth are useful for many purposes: crop evaluation, water pollution, and other studies of the environment. With the developing ability to produce color pictures and data from other wavelengths, such as various multispectral bands in the infrared range, satellite observations of our global environment may become increasingly important. The quantity of data already available from satellite observation is more than can be fully processed manually on a routine basis. Hence, the need to develop automatic methods of processing satellite data is becoming more urgent.

The goal of removing the need for human visual processing of satellite data is the general area in which our work lies. Automating human visual tasks by means of a computer may never be fully achieved; but, if we limit the computer tasks to deriving cloud vectors automatically, we are confident of some success on the simpler cloud motions. By "success," we mean automatically producing enough cloud motion vectors that are:

- Sufficiently accurate,
- In reasonable limits with respect to computer processing time and costs, and
- Compatible with the scale of operations established by current numerical prediction models.

We will discuss in more detail a systems approach to developing the automatic programs. This approach does not merely involve a programmer writing algorithms for the computer, but it also makes extensive use of a high-performance display, the SRI cloud-tracking console, and discussions with many of the researchers at SRI in this field of work.

II OBJECTIVES

The overall objectives of this research project were stated in the proposal, calling for enhancement and generalization of the SRI automatic cloud motion computation method as follows:

(1) To apply to multilayered cloud patterns.

(2) To experiment with the multilayered cases, using a man-machine interactive computer system.

(3) To investigate the use of infrared data in a preliminary manner.

In more detail, the specific objectives of this research are:

(1) To generate artificial data representing clouds (clusters of data points) and display these data on a computer-driven CRT display screen for visual inspection. (The SRI PDP 11 and Vector General display is suitable for this function.)

(2) To simulate multilayered cloud motion fields in the above manner. For example, the Inter Tropical Convergence Zone (ITCZ) has a characteristic motion field.

(3) To convert and interface the data generated on the PDP 11 and display, with the CDC 6400 on which the ISODATA* programs can be run. These programs can automatically detect some simple motions of clouds in sequences of picture-tapes from geosynchronous satellites, such as the ATS-III. Therefore, the programs on the CDC 6400 must be modified to accept artificial data, which are not constrained to the same gridded picture format as the picture-tapes.

---

*ISODATA is a general identifier for the group of computer programs we are working on for cloud tracking. In a section of this report we will later discuss all the programs in this group, individually.

Preceding page blank

(4) To simulate infrared (IR) data to accompany the visible data, so that cloud height determination can be accomplished by ISODATA automatic motion detection programs.

(5) To modify the ISODATA programs to accommodate the fourth variable (IR), so that differing cloud motions at various heights (IR temperatures) can be discriminated.

(6) To rank the motion vectors in terms of a fitting factor, and to adjust the contributions of the fitting factor components so that our intuitive ranking agrees with the computed ranking.

(7) To submit photographs of the artificial clouds to be tracked on the manual cloud-tracking console at SRI for human verification of the computer-detected motions.

III THEORY OF AUTOMATIC CLOUD TRACKING

One of our most important assumptions in cloud tracking theory is that the human cloud tracker can perform more effectively than a computer if there is no time limit or data volume restriction. Therefore, we discuss and solicit comments on our computer results from those individuals who regularly use cloud-tracking consoles and have a sound meteorological background. Thus, comparison between results produced by man and machine is an important input to evolving our algorithms. On a related project (under NASA Contract NAS 5-21776),[5] we have carried out a detailed comparison of the cloud vectors produced by the "ISODATA" method and the cloud console vectors produced by humans. The "ISODATA" method is the collection of computer programs that we use to perform automatic cloud tracking. Note that ISODATA is the name of the clustering algorithm that makes up one of the programs in the system. To avoid confusion, we will use the term ISODATA method to denote the entire system and ISODATA program to denote the particular implementation of the clustering algorithm.

A. Axioms of Cloud Tracking Theory

The following maxims or axioms are the basis of our current tentative theory regarding cloud tracking. These axioms may appear obvious and not significantly detailed or specific to be of use in solving some of the programming details that must be produced as an end result. However, they are of overriding scope and must always be remembered as assumptions, even when working on the programming details. We propose them here, not in a sense of finality, but as a starting point for their refinement, so that we have something specific in the way of axioms to build upon.

- Axiom 1--Man is the best cloud tracking perceptual "mechanism"; however, the man/machine combination is the best cloud tracking system. Even in the simplest systems, man is usually augmented by some machinery, e.g., a camera, pencil and paper, a movie projector, satellite, CRT display, cloud console, computer, film loop, or digitizing board.
- Axiom 2--Human cloud trackers can be significantly enhanced by means of time-lapse display equipment (such as the SRI cloud console).
- Axiom 3--Automatic tracking functions can replace human tracking functions when human trackers can explain what they are doing and when these explanations can be programmed.
- Axiom 4--Because of the limitations of programming languages even some operations that can be described clearly by cloud trackers cannot be programmed. (The limitations of programming languages are partly caused by computer hardware limitations also, but in these applications we are more concerned with software, since hardware developments are not part of this project.)
- Axiom 5--Clouds can only be tracked if they persist from one frame to the next.
- Axiom 6--For clouds to be seen as persisting, they must be recognized from one frame to the next--hence, we understand the importance of characterizing pairs of clouds by statistical (clustering) descriptors (e.g., by using the ISODATA clustering program, defined in more detail later) and combining differences between these descriptors into a single recognition number called the fitting factor for two clouds. Two clouds are said to be the same (recognized) cloud if their fitting factor is sufficiently low. (Actually the number should be called the misfitting factor because the larger it is, the less likely the two clouds will fit. We discuss the fitting factor in more detail later.) Here, fitting is a pairing function, though we would like to extend the "pairing" over more than two frames.
- Axiom 7--The application of the ISODATA method to picture data is a process of dynamic clustering. In previous reports we have described how the ISODATA and MOTION programs have been used: first, the ISODATA program detects brightness centers as centers of clusters, and then MOTION detects their vector motions by comparing successive clusterings. We now view this dual process (ISODATA then MOTION - the

ISODATA method), as a single process of dynamic clustering, or clustering-in-motion. The dynamic tracking of visual objects has absorbed much study in the field of pattern recognition, animal and insect biology, and human vision. We feel that the ISODATA method represents fairly well (at this stage of research) what the human cloud tracker is doing when he follows a set of cloud elements through a series of several time-lapse pictures.

- Axiom 8--Cloud motions tend to be continuous. (We call this the continuity principle.) Visual cloud tracking makes strong use of continuity in time and space. Therefore, the program should only accept vectors (i.e., join sequential cluster centers) if:

  (1) Time continuity is shown by the vectors persisting over several frames.

  (2) Space continuity is shown by adjacent* vectors having a similar speed and direction. (Currently, we compare each vector to the average vector for the whole region. This may be too global a comparison, but is computationally simple.)

  The MOTION program (defined in more detail later) tests the connections between pairs of points in successive picture frames, and rejects the connections if the principle of continuity is violated. The vector filtering program rejects vectors that do not persist in at least two motion vector frames (i.e., three picture frames)--thus, testing the connections between triples of points in successive frames.

B. Definitions of ISODATA and MOTION Programs

We give here a definition of two computer programs frequently referred to in this report. ISODATA was originally considered an acronym for Iterative Self-Organizing Data Analysis Technique. The first ISODATA computer programs were written in ALGOL, in 1966.[6] Since that time, many improvements and special-purpose versions (mostly in FORTRAN) have

---

*Adjacency must include the height of clouds because different height layers often move in opposite directions.

been made and the subsequent programs have been applied to a wide range of data. The purpose of all ISODATA programs is to perform clustering of data.

Clustering is a process of forming a condensed version of a set of data, so that the essence of an extensive and complex data set is represented by a new value called the cluster center. The cluster center is usually the mean value, or average; it is accompanied by other statistical descriptors, such as standard deviations, number of points in each cluster, etc. When clustering cloud data, we seek to describe the considerable detail of satellite photographs by suitable clusters, so that clouds can be accurately and briefly characterized.

The MOTION program grew out of a need to perform successive clusterings in time sequence and characterize the changes occurring in the data. Our first program to perform dynamic clustering was written by Roy Endlich.[3] The MOTION program uses two sequential sets of cluster centers as inputs, and finds correspondences between cluster centers based upon their associated statistical descriptors. A pair of clusters is recognized (i.e., paired) if they match each other closely enough in terms of a fitting factor. This fitting factor is currently a heuristic combination of statistical descriptors given by the following formula:

$$FF = \left[W_x \Delta X^2 + W_y \Delta Y^2 + W_B \Delta B^2 + W_{IR} \Delta IR^2 + W_n \left(\frac{\Delta n^2}{n_1 + n_2}\right) + W_\sigma \left(\Delta\sigma_X^2 + \Delta\sigma_Y^2 + \Delta\sigma_B^2 + \Delta\sigma_{IR}^2\right)\right]^2 + W_c \Delta C$$

where

- FF is the fit factor value. The W's are weights for the respective factors; and $W_x = W_y = W_B = 1$, $W_{IR} = 2$, $W_n = 25$, $W_\sigma = 2.5$ and $W_c = 3$.
- Δ stands for the difference in the respective statistical descriptors of a pair of cluster centers.
- X is the measurement, in arbitrary units (1 to 120), in the east-west direction. Y is the measurement, in arbitrary units (1 to 70), in the north-south direction. B is the average brightness of a cluster (0 to 255). IR is the average infrared temperature of a cluster (160 to 330 degrees Kelvin). n is the number of points in a cluster, in particular, n_1 is the number of points in the cluster at time 1, and n_2 is the number of points in the cluster at time 2.
- The σ's are the standard deviations of the various quantities. Thus, σ_x is the standard deviation in the east-west direction, σ_y is the standard deviation in the north-south direction, and so forth.
- The $W_c \Delta C$ term makes use of the assigned cluster number. If the assigned cluster numbers for the two clusters is the same, $\Delta C = 0$; if the cluster numbers are different, $\Delta C = 1$.

IV EXPERIMENTAL METHOD

The method of conducting our experiments has been to generate artificial data (representing clouds) and then to process these data using the dynamic clustering programs (ISODATA and MOTION) to see if they can recover the same information from the data as we used to generate it. This method is sound because it is complete, having a closed-loop form.[4] Discrepancies that arise between the two sources of information, at the input and output of the closed loop, denote errors in the dynamic clustering program. If the errors are statistically significant, modifications to the program are necessary. Usually the errors are of a gross kind, but the search for a remedy may require introspection concerning the nature of the human visual tracking process and how to describe this within the practical limitations of the computer system. The commands shown in Table 1 are implemented in the system, constituting the data-generating language. The language can be input either from the on-line Teletype on the PDP 11 or from the card reader on the PDP 11. In addition, the same cards may be fed into the CDC 6400, which operates only in a batch mode. Each command starts with a single character and may be followed by numerical parameters. This is a very simple language, but it is sufficient to allow great conveniences in practical use.

Specific modifications to the data-generating language for this work were to add the commands "B" and "A." These allow for the specification of the brightness (B) and altitude (A) of the cluster. The brightness is given on the digital satellite tapes by means of an 8-bit code, thus specifying the light intensity in the range 0 to 255. The inclusion of infrared data in the objective system, by means of the command "A," is

Preceding page blank

Table 1

COMPUTER COMMANDS FOR THE CLUSTERING LANGUAGE

Number	Commands (operator options)	Default Values
1	A Altitude of Cluster I	
2	B Brightness of Cluster I	
3	C Cluster center for Cluster I	(0,0)
4	D Standard deviation for Cluster I	(300,300)
5	E Exit to monitor	
6	F	
7	G Go-start execution--generate new cluster centers	
8	H Use old cluster centers (assumes no changes to specs)	
9	I Cluster center index	1
10	J	
11	K	
12	L List value of parameters of Cluster I on teletype	
13	M	
14	N Number of cluster centers	1
15	O List of commands on line printer (LP)	
16	P Number of points in Cluster I	100
17	Q Debug print if $Q \neq 0$	0
18	R Initial random number value--two-integer values	
19	S Switch from card reader to teletype; or vice versa	Teletype input
20	T	
21	U	
22	V	
23	W Rewind data file	Start of file
24	X Transform coordinate scale factors for X and Y	(120, 70)
25	Y	
26	Z Zero summary statistics	

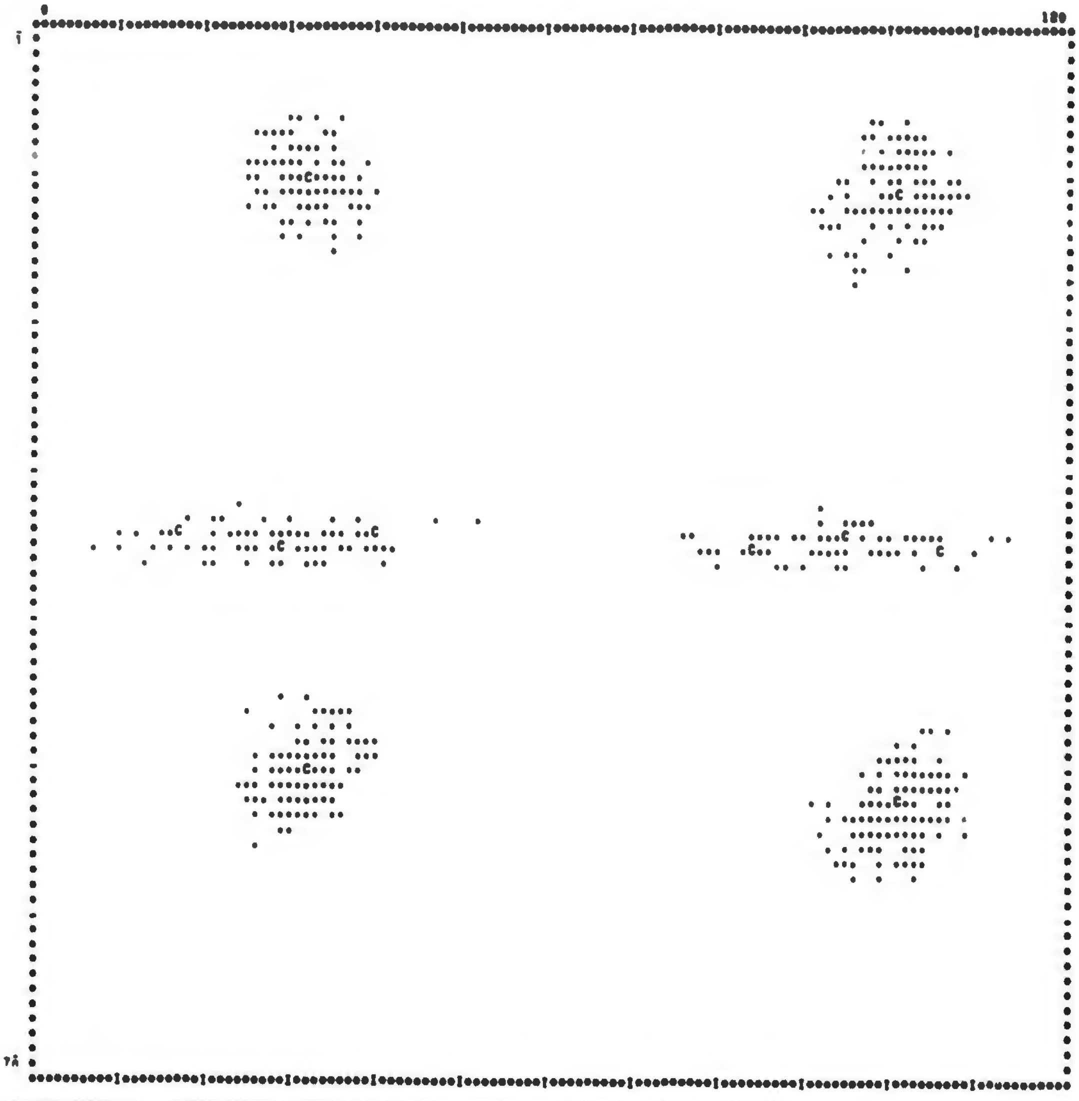
recognized as an important extension, since different cloud layers can be discriminated in this way. Such data are now becoming more readily available from satellites; they can be realistically simulated by reference to certain cases viewed in the visible spectrum by ATS-III, and in the infrared spectrum by Nimbus-4, NOAA-2, and NOAA-4 (launched 8 November 1973). Preliminary tests of the application of simulated infrared data have been made using the interactive computer system.

We now illustrate the generation of six artificial clouds and their subsequent clustering and tracking by means of the dynamic clustering program. This example displays the essence of the approach taken. Most of our later experiments are merely elaborations on this basic theme, concentrating on various particular aspects of the programs where it is "experiencing" difficulties, i.e., making errors.

By means of the cluster-generating language, the data for this example were generated and displayed on the CRT screen of the PDP 11, and also, as shown in Figure 1, were plotted on the CDC 6400 line printer. Some remarks concerning the interpretation of such printer plots is given next.

A. Interpretation of Line-Printer Plots

When inspecting the printer plots of the data, each mark, ".", may represent several data points. This occurs because the printer plot is the result of a data transformation from the ±2047 unit scale (in horizontal and vertical) of the CRT to the 120 by 70 scale suitable for line printer output. (See Figure 8, in Section V of the report, for a photograph of the displays in which each data point is displayed separately and more precisely.) Such a quantization error also occurs in the MOTION program output printer plots. For these reasons, the plots are usually accompanied



```
CONVERGED AFTER  3 ITERATIONS
```

FIGURE 1 A PRINTER PLOT FOR CLOUDS OF ARTIFICIAL DATA by more accurate tabular outputs from the computer. When it appears that a MOTION vector is in error or if a cluster center (shown by a "C" in the plots) appears to be slightly off-center, these tables should be consulted for verification.

B. Example of Experimental Methods

For our simplest case we generated four artificial clouds that were placed near the corners of the display-screen rectangle and gave them a uniform motion. After we checked that the program could track these, another four clouds were added very close to the first set, overlapping so that eight clusters were called for by the generating language; although to the eye (and to the ISODATA program), only four clouds appeared to be generated. This overlapping technique allows us to move the separate components of each double cluster independently, allowing more extensive--but controlled--shape changes. In addition, a second, high-altitude layer of clouds was generated and was moved in an almost opposite direction to the low-level clouds. This simulates a condition we have observed with real satellite data. The printer plot for the first picture in the sequence is shown in Figure 1. These data were generated using the computer commands given in Table 2. Note that 10 clusters are called for ("N10"). Each of the four round corner groups consists of two clusters, and the two central, elongated clusters are generated from a single cluster-specification each. In this latter case, the horizontal deviation value is specified as six times larger than the vertical deviation (i.e., D 300 50).

The second picture of these clouds is shown in Figure 2, and by visual comparison with Figure 1 it can be seen that the round clouds in the corners have moved down diagonally towards the lower left corner of the figure, and the elongated clouds near the center of the figure

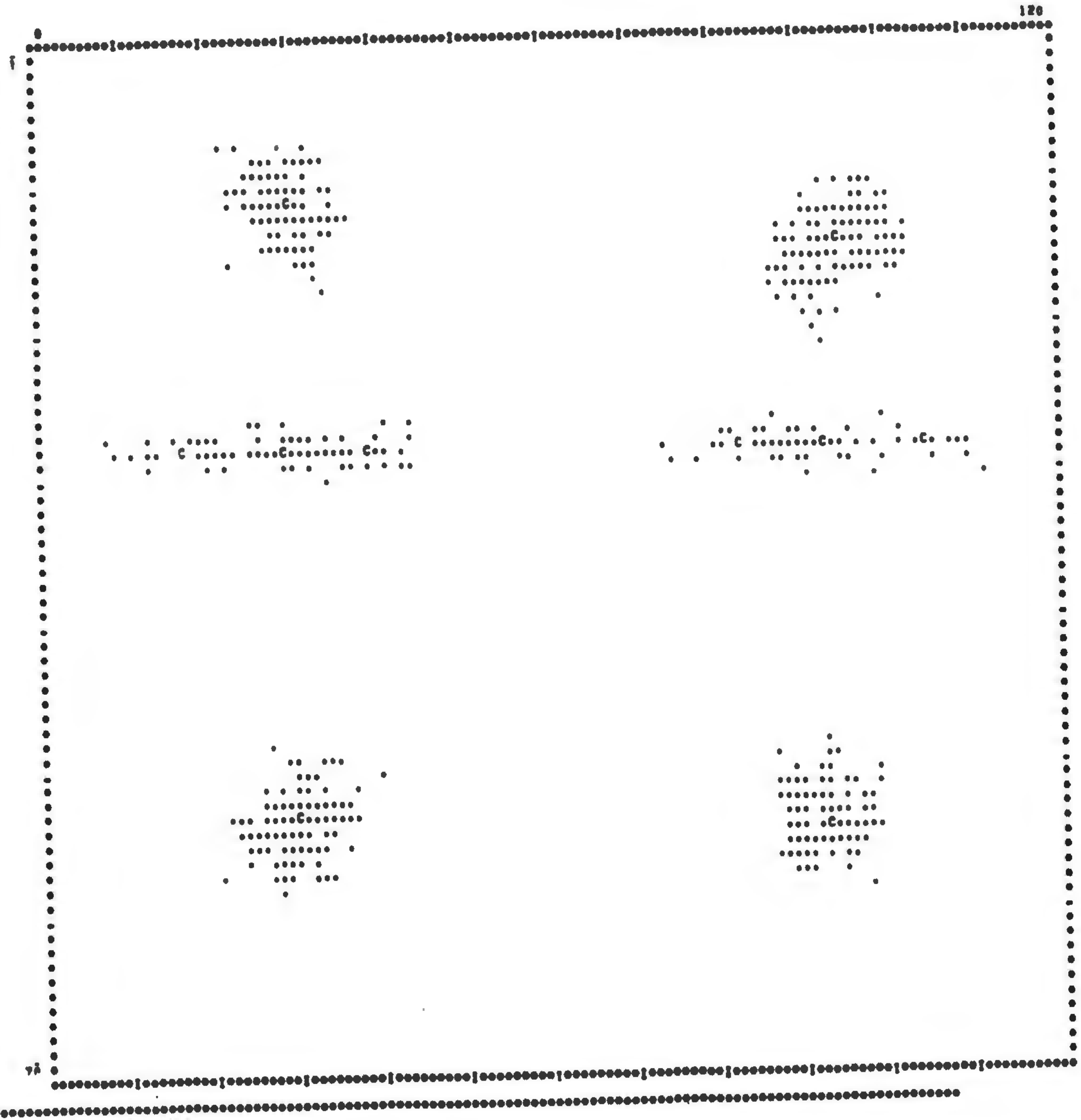
Table 2

COMPUTER COMMANDS FOR GENERATING
10 CLUSTERS OF ARTIFICIAL DATA

```
N10
   I1 P100 C -1000 1500 D 100 100
   I2 P100 C  1500 1500 D 100 100
   I3 P100 C -1000 -1000 D 100 100
   I4 P100 C  1500 -1000 D 100 100
    I5 P 40 D 100 100 C  -900 1400
    I6 P 45 D 100 100 C 1300 1300
    I7 P 55 D 100 100 C -850 -850
    I8 P 60 D 100 100 C 1350 -1150
    I9 P 80 D 300 50 C 1100 0
    I 10 P 80 D 300 50 C -1100 0
                                        G
```

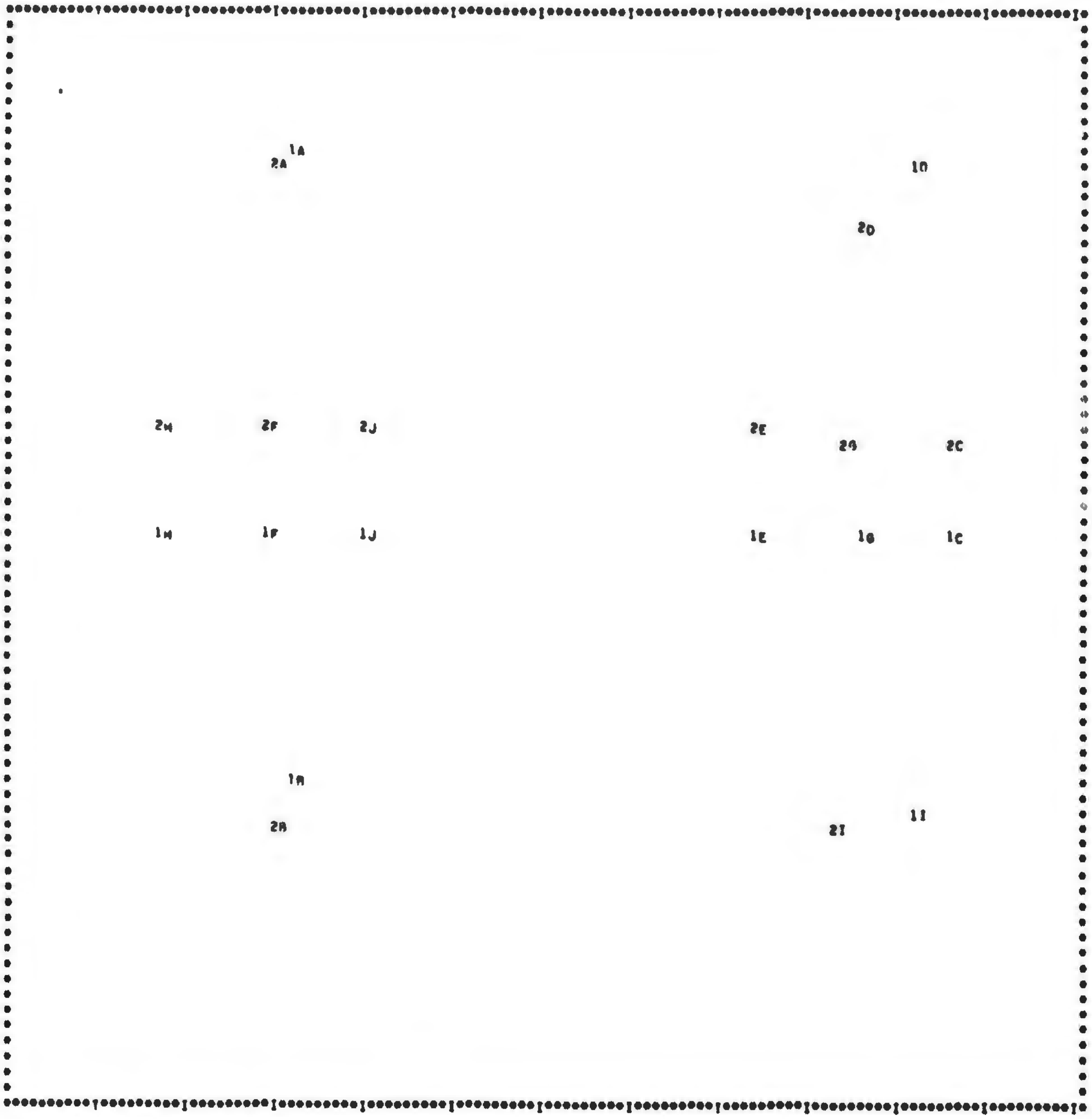
have moved vertically upwards. Note that although the general shapes of the clouds are similar in the two pictures, the dot patterns are quite different. This statistical similarity is achieved by using the same statistical parameters, but without resetting the random number generator to the same initial condition at the start of the data generation process for each picture. An example in which we do reset the random number generator is given later.

The ISODATA and MOTION programs were able to provide suitable vectors for this simple two-layer case. The output of the MOTION program is given in Figure 3, which shows pairs of cluster centers labeled from A to J, and prefixed by a "1" or "2," denoting picture sequence. In interpreting these printer plots of motion vectors, each pair is labeled with a letter of the alphabet (or other single character after the whole alphabet has been used). If no single character identifier is given alongside the "1" or "2," this means that no pairing was made by the



```
CONVERGED AFTER  2 ITERATIONS
```

FIGURE 2 PRINTER PLOT FOR THE SUCCESSIVE PICTURE OF ARTIFICIAL DATA



```
AVERAGE U AVERAGE V  NO.VECTORS

   -1.97       4.18       7.00
```

FIGURE 3 CLOUD MOTION VECTORS PLOTTED ON A LINE PRINTER program. "1" and "2" denote the positions of the cluster centers in the first and second pictures. These vectors are the end product of our cloud tracking system and are suitable input data for a numerical prediction model.

We eventually plan to add further complexity and reality to our simulation by allowing the high-level clouds to pass over the low-level clouds--a condition we carefully excluded from the case described above.

Some of the more detailed program features are as follows:

(1) A multiple execution approach has been provided with respect to the processing of cloud-vector data. All vectors which were discarded by execution of MOTION on the original data have been entered on a new file created in exactly the format acceptable as input to MOTION. By use of system control cards (REWIND and a second set of execution cards) the file may be used for another execution of MOTION. The purpose of this is to allow for detection of a second set of vectors that may have a mean vector vastly different from the first mean vector. This case of two different mean vectors would typically be found in multilayer cloud tracking examples.

(2) Descriptive outputs provided are:

  (a) All input clusters generated have been tagged with an associated letter which is printed in the first print-plot to aid in identifying the various cluster centers and their associated data visually.

  (b) A print-plot has been provided along with associated mean vector, fitting factor, etc. This print-plot creates a visual representation of only those vectors with matching cluster center number, as generated in the ISODATA program.

  (c) "Edge" cluster centers have been marked with a "-" to the right of the cluster center letter in all print-plots after the first.

  (d) For the data after each normal print-plot, letters are associated with those vectors which have been selected for further consideration. Those vectors which have been discarded have no such letters.

(3) The program is able to complete correct processing regardless of whether the data being used is from an artificially generated source or from ATS data.

(4) A weighting factor has been provided for the fitting factor which accentuates the pairing of cluster centers with the same generated cluster number.

Many of these features will be displayed in results reported in the following section. This section has been intended to give the reader an understanding of our experimental method before discussing more complex examples.

V EXPERIMENTS USING PDP 11 AND CDC 6400 COMPUTERS

Various experiments have been undertaken in attempts to improve some aspects of cloud tracking performance using objective methods. A simple example was given in the previous section to illustrate our experimental method. Now we will discuss, in considerable detail, the tracking of clouds in an interesting motion field.

A. The Intertropical Convergence Zone

We generated an artificially specified cloud motion field that characterizes the usual cloud motions in the Intertropical Convergence Zone (ITCZ). This motion field is shown in Figure 4. It is derived by plotting the generating language data for the positions of correspondingly numbered cluster centers in three successive frames. Later, we will compare this objective cloud motion field with that produced by human visual tracking using the SRI cloud console, and by the computer algorithm for dynamic clustering.

A way to compare the effectiveness and accuracy of the clusterings individually is to print out the clusters found by the ISODATA program in a format compatible with the input data generating language. This was done for the SRI project to develop new pattern-recognition methods,[4] but is not so useful here because we are more interested in the recovery of characteristic vectors rather than the recovery of the original clusters in computer language form. Hence, we compare the input and output motion vector fields here. In simulating so complex a motion field as the ITCZ, the cloud data are neither very separated nor highly clustered,

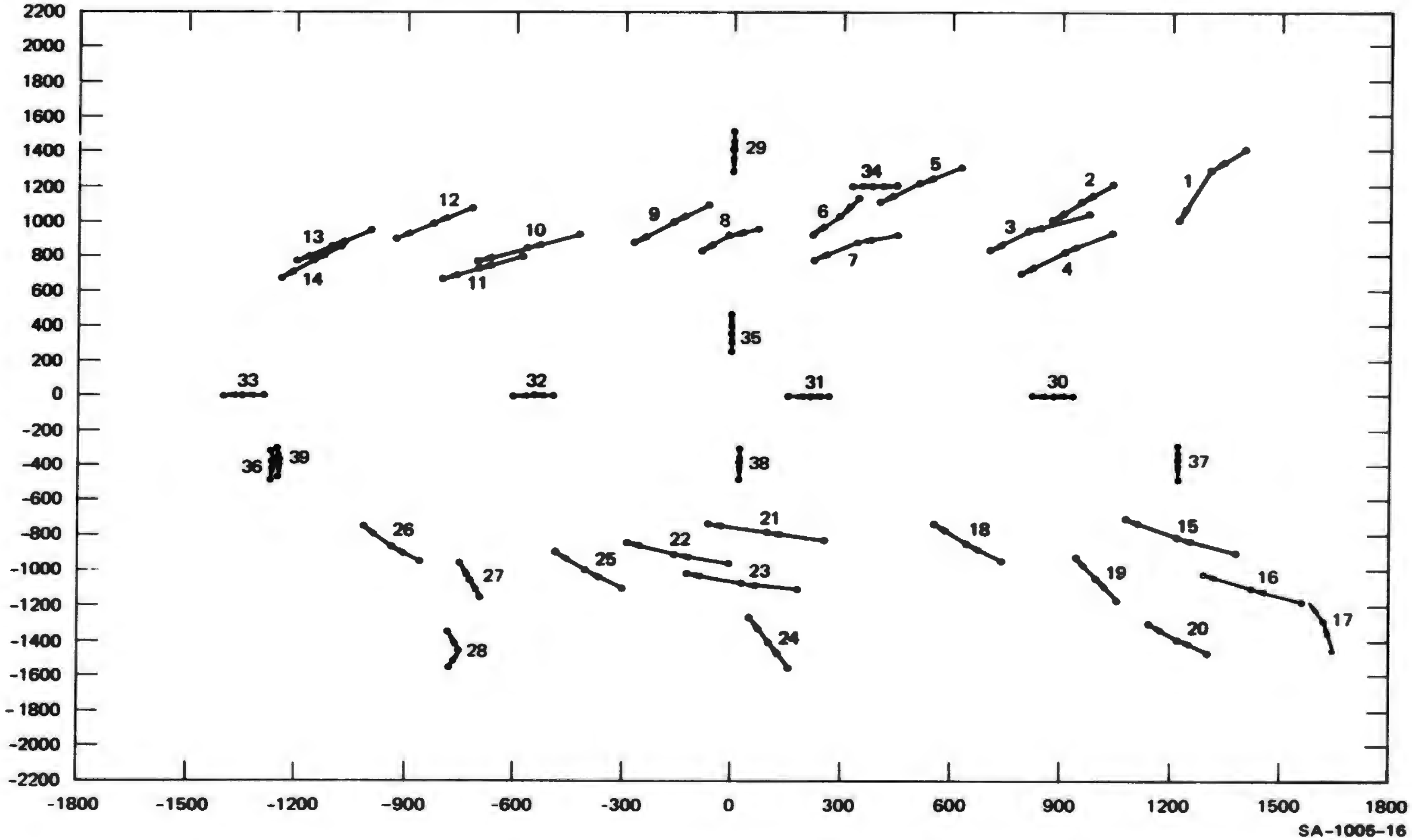


FIGURE 4 MOTION VECTORS DERIVED BY PLOTTING THE GENERATING-LANGUAGE DATA

so we cannot expect to retrieve the identical clusters that are fed into the recognition system as inputs. However, we must retrieve the general characteristic motions if the dynamic clustering is to be successful.

B. Three Successive Artificial Pictures of the ITCZ

In this simulated version of the ITCZ, high-level cloud data are included. In the data-generating language, 39 clusters are called for, as shown in Table 3. Figure 5 illustrates the data in computer printout form. The second picture in the sequence is shown in Figure 6, where the cluster center locations on all 39 clusters were modified. However, note from the data-generating language given in Table 4 that only the changed values in each cluster need be given and all values not changed are merely repeated as before. Note particularly that the statement on the second line, "R 5103 2187," sets the random number generator to particular starting values. These are the values used at the beginning of each picture, so that identical pseudo-random numbers are used in the same sequence for each of the 998 points in the data set. This produces very similar detailed structure in each picture.

The final picture in this sequence is shown in Figure 7, again involving new positions along a plausible trajectory for all 39 clusters.

Besides displaying the data in print-plot form, a more accurate data plot can be photographed from the PDP 11 display screen as shown in Figure 8, which shows the three pictures in the sequence corresponding to the plots of Figures 5, 6, and 7. The three pictures have been photographed and the negatives were used as input to the SRI cloud console. Viewing of the time-lapse display gives a definite impression of motion, and the experienced cloud-tracking operators say that the cloud motions in these pictures are easily trackable.

Table 3

COMPUTER COMMANDS FOR GENERATING 39 CLUSTERS SIMULATING THE ITCZ

```
N 39
  I 1 P 14 C 1400 1400 D 50 50 B 65 A 250
  I 2 P 11 C 1040 1200 D 50 60 B 60 A 245
  I 3 P 19 C 930 1050 D 50 40 B 63 A 243
  I 4 P  9 C 1040 920 D 40 50 B 60 A 240
  I 5 P 18 C 620 1300 D 50 50 B 85 A 255
  I 6 P 20 C 340 1125 D 60 30 B 75 A 250
  I 7 P 19 C 450 935 D 40 40 B 64 A 247
  I 8 P 10 C 60 950 D 45 50 B 80 A 253
  I 9 P 23 C -65 1085 D 70 65 B 72 A 245
  I 10 P 17 C -425 930 D 50 45 B 70 A 253
  I 11 P 10 C -585 785 D 45 45 B 67 A 245
  I 12 P 8 C -720 1070 D 30 40 B 85 A 246
  I 13 P 16 C -1000 930 D 45 55 B 90 A 253
  I 14 P  9 C -1075 860 D 55 45 B 78 A 246
  I 15 P 14 C 1375 -900 D 50 50 B 75 A 247
  I 16 P 11 C 1565 -1190 D 50 60 B 78 A 244
  I 17 P  9 C 1645 -1445 D 50 40 B 82 A 250
  I 18 P 16 C 740 -950 D 60 30 B 62 A 248
  I 19 P 13 C 1050 -1160 D 50 40 B 66 A 245
  I 20 P 11 C 1300 -1465 D 45 55 B 60 A 242
  I 21 P 10 C 255 -820 D 50 35 B 60 A 247
  I 22 P 23 C 0 -955 D 40 45 B 67 A 252
  I 23 P 23 C 200 -1110 D 45 55 B 78 A 255
  I 24 P 18 C 160 -1555 D 30 30 B 85 A 257
  I 25 P  8 C -300 -1100 D 25 25 B 78 A 244
  I 26 P 25 C -860 -960 D 40 40 B 67 A 243
  I 27 P 24 C -700 -1145 D 30 35 B 77 A 247
  I 28 P 29 C -775 -1525 D 25 25 B 87 A 257
  I 29 P 55 C 1565 0 D 200 50 B 150 A 239
  I 30 P 35 C 925 0 D 100 50 B 120 A 230
  I 31 P 60 C 255 0 D 200 50 B 150 A 235
  I 32 P 65 C -505 0 D 250 50 B 160 A 238
  I 33 P 55 C -1295 0 D 200 50 B 155 A 233
  I 34 P 43 C 1200 330 D 225 75 B 45 A 220
  I 35 P 63 C 0 350 D 375 75 B 43 A 223
  I 36 P 39 C -1265 340 D 275 75 B 41 A 225
  I 37 P 38 C 1220 -330 D 250 75 B 44 A 222
  I 38 P 66 C 20 -350 D 300 75 B 45 A 220
  I 39 P 43 C -1250 -340 D 325 75 B 42 A 224
 G
```

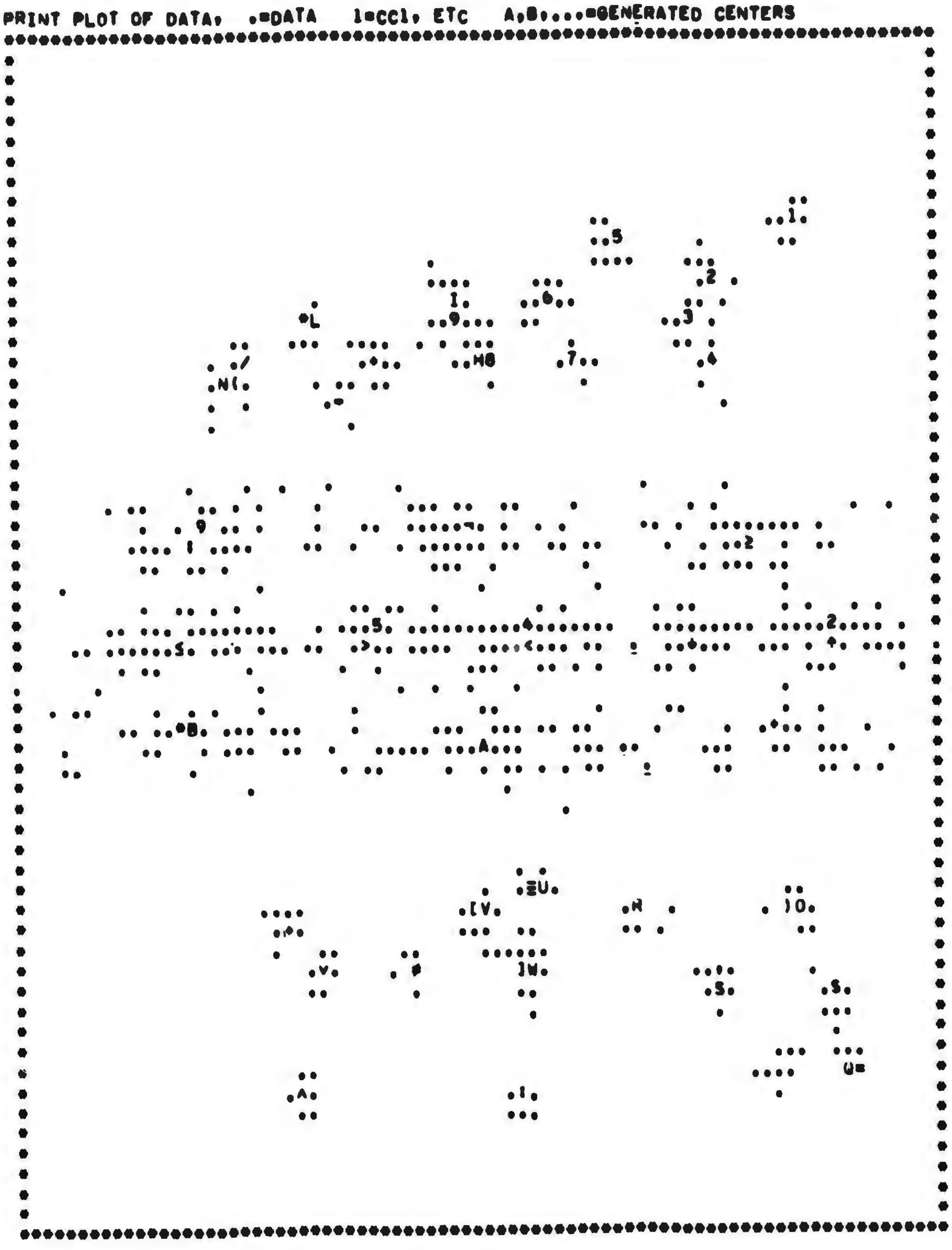


FIGURE 5 39 CLUSTERS OF THE ITCZ AS GENERATED FROM THE DATA OF TABLE 2

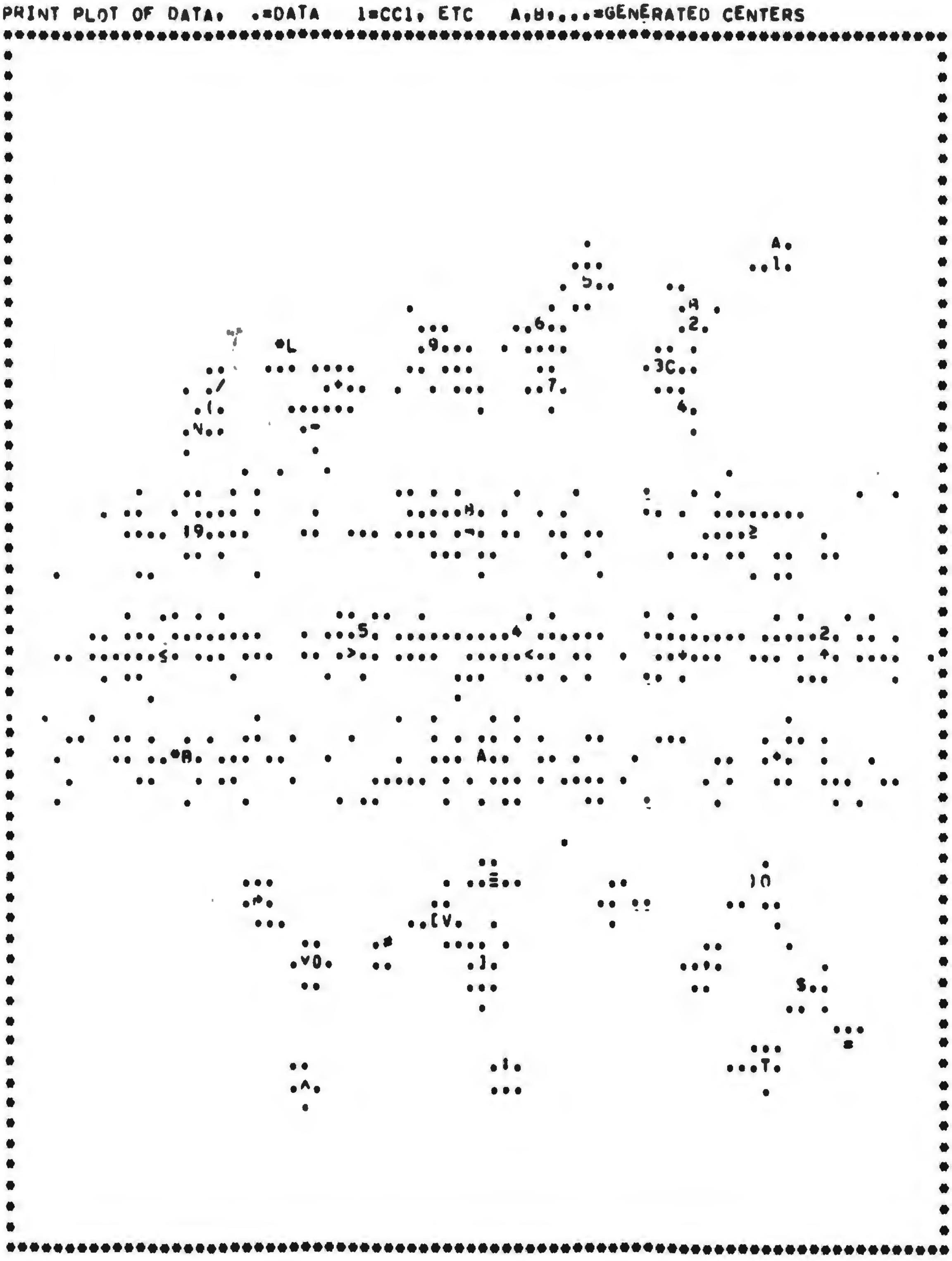


FIGURE 6 THE SECOND PICTURE IN THE SEQUENCE FOR THE ITCZ

Table 4

THE DATA FOR GENERATING PICTURE 2

```
N 39
 R 5103 2187
  I 1 C 1310 1295
  I 2 C 941 1090
  I 3 C 807 947
  I 4 C 900 810
  I 5 C 500 1210
  I 6 C 280 1035
  I 7 C 325 865
  I 8 C -12 902
  I 9 C -175 980
  I 10 C -575 855
  I 11 C -705 730
  I 12 C -825 990
  I 13 C -1105 855
  I 14 C -1165 767
  I 15 C 1215 -810
  I 16 C 1425 -1110
  I 17 C 1625 -1300
  I 18 C 640 -850
  I 19 C 1000 -1050
  I 20 C 1215 -1390
  I 21 C 80 -780
  I 22 C -160 -900
  I 23 C 20 -1065
  I 24 C 110 -1425
  I 25 C -400 -1000
  I 26 C -940 -860
  I 27 C -720 -1045
  I 28 C -750 -1450
    29 C 1515 0
  I 30 C 875 0
  I 31 C 205 0
  I 32 C -555 0
  I 33 C -1345 0
  I 34 C 1200 380
  I 35 C 0 400
  I 36 C -1265 390
  I 37 C 1220 -380
  I 38 C 20 -400
  I 39 C -1250 -390
 G
```

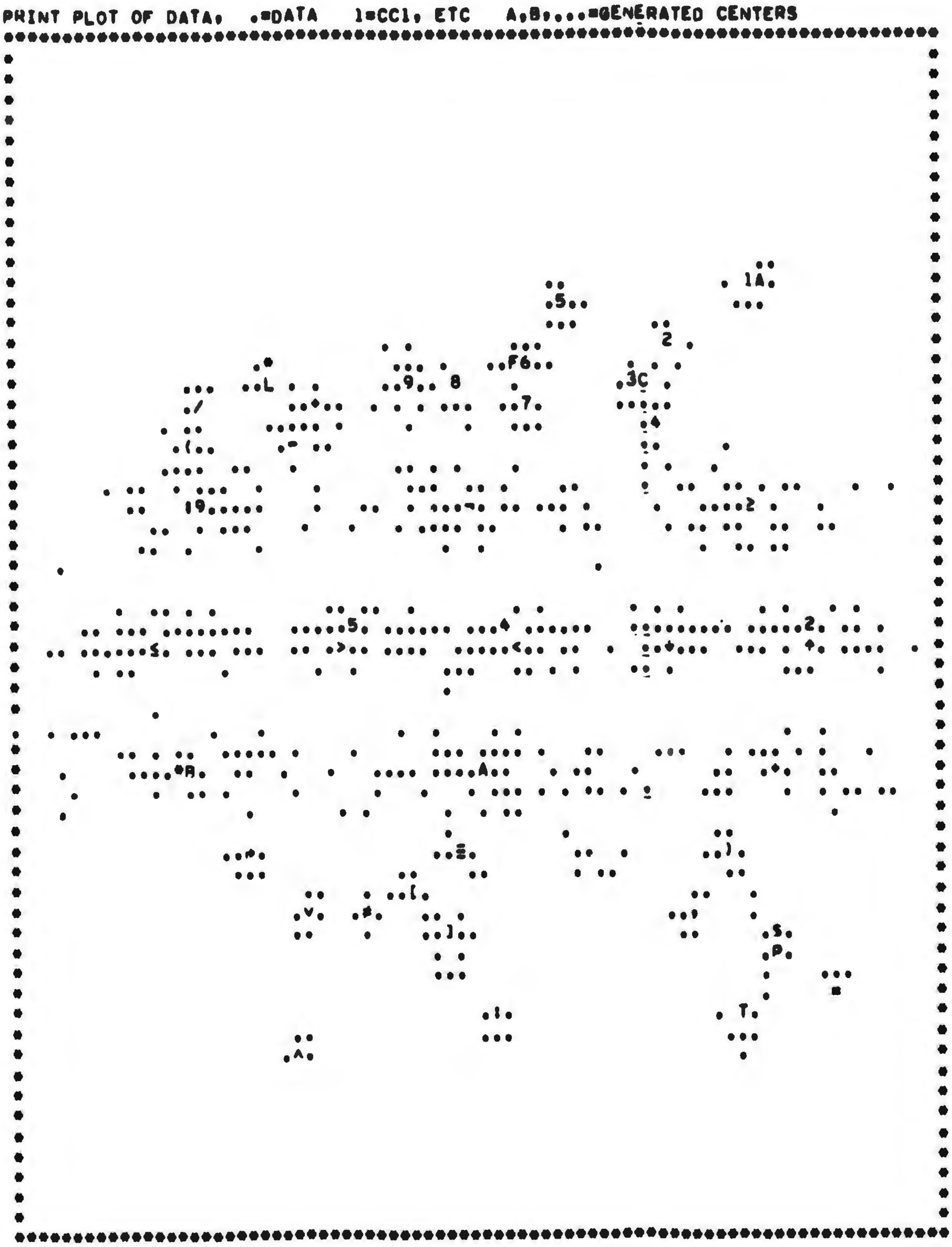
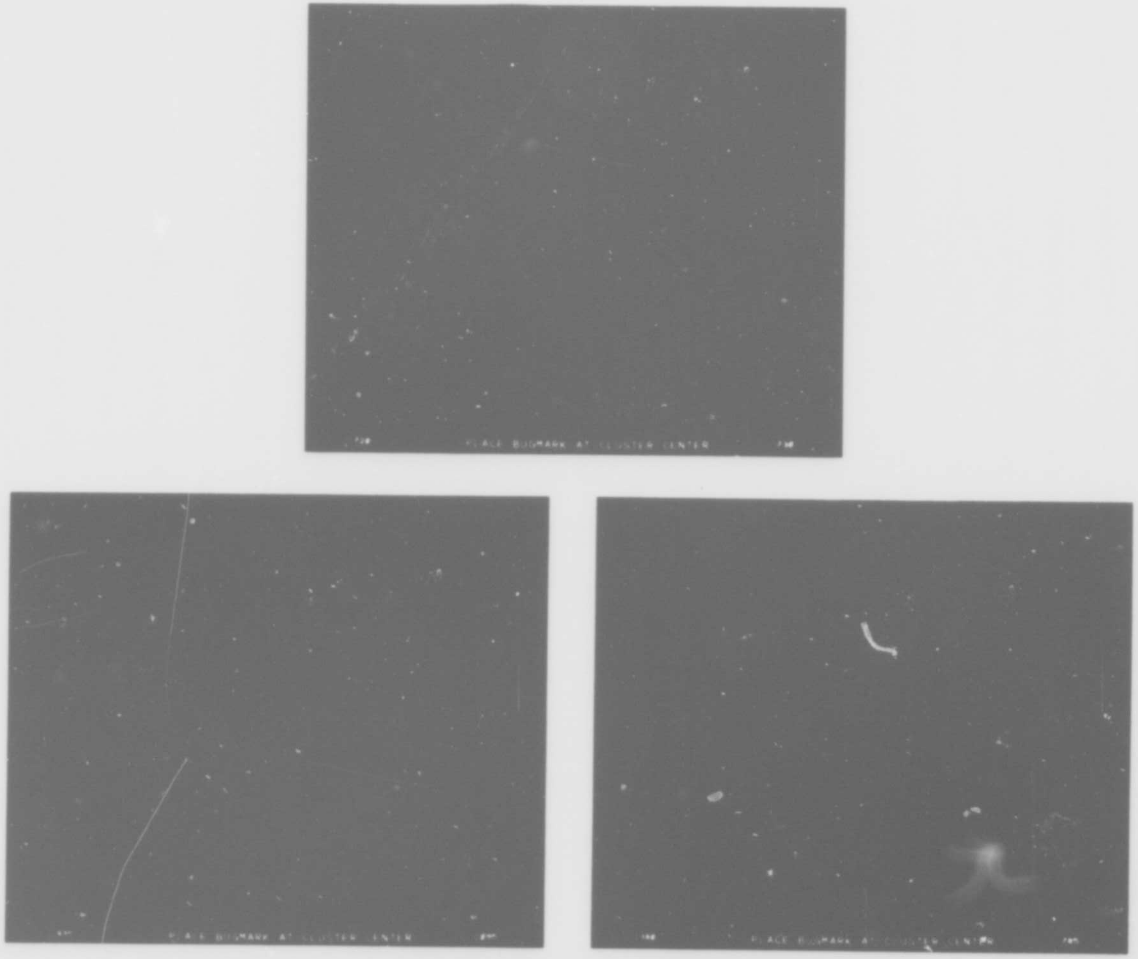


FIGURE 7 THE THIRD PICTURE IN THE SEQUENCE FOR THE ITCZ



SA-1005-17

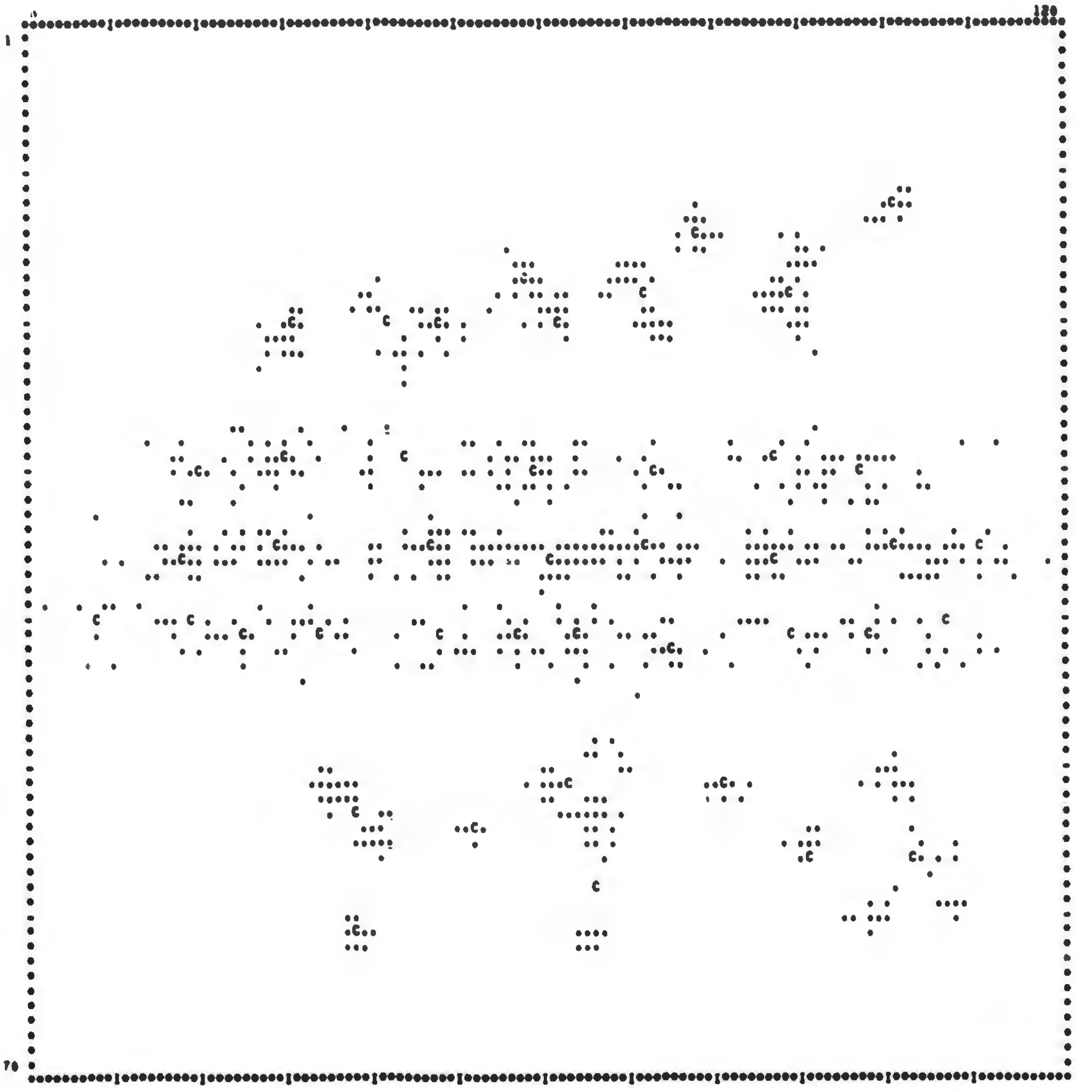
FIGURE 8 PHOTOGRAPH OF THE PDP-11-CRT SCREEN DURING DISPLAY OF ITCZ DATA

The next step in the computer program clusters the data, and convergence for the first picture was achieved in three iterations, producing 43 clusters. The second picture clustering converged in three iterations also, but 44 clusters were produced. The third picture converged in two iterations, also with 44 clusters. Note that although 39 clusters were specified in the data-generating language, the ISODATA program found 43 or 44 clusters. This difference shows that the generating and clustering mechanisms are not identical in detail. However, inspection of the cluster center plot superimposed on the data, as given in Figure 9, shows that the centers are not unreasonably placed according to human visual inspection. We cannot claim that this placement is optimum or unique in an exhaustive sense, but it is stable because of its convergence.

ISODATA clustering is applied to each of the three pictures in sequence to produce three sets of cluster centers. Next, these cluster centers must be linked in time to produce the motion vectors using the MOTION program.

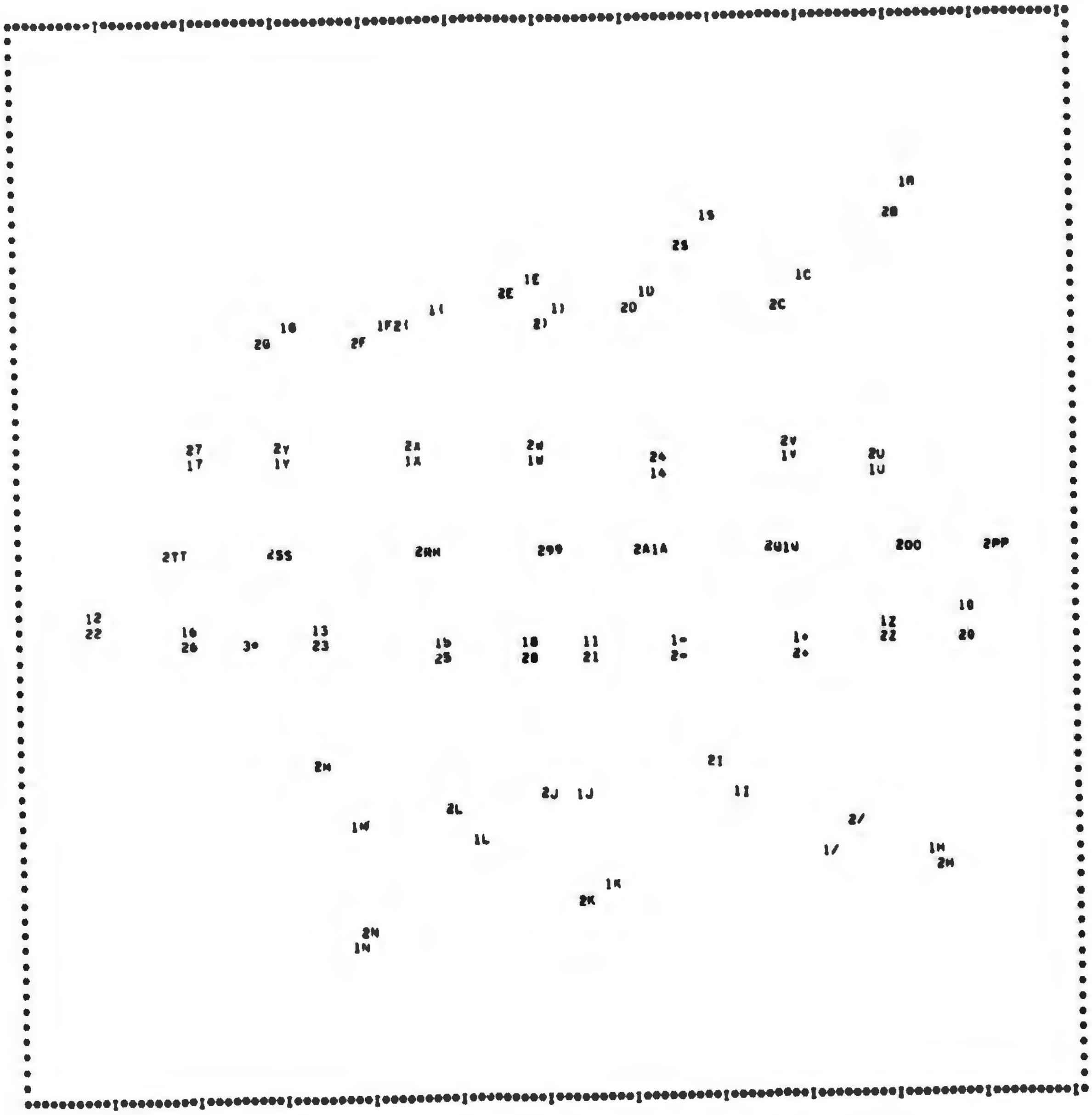
At one time in the experiment, the ISODATA cluster center numbers were used to pair or fit successive clusters. An example of this is shown in Figure 10. From this it can be seen that this use of cluster number for pairing provides a good way to start the pairing process.

Figure 11 shows the situation after four iterations of the MOTION program in which the single-character identifiers on the cluster pairs are very different than in Figure 10. However, the differences in the pairings are rather minor. The relabeling of the pairings takes place because these pairings are ranked according to their probability of fit with each other in accordance with several criteria that contribute to the overall fitting factor. This result confirms what we might expect, a priori: the distribution of the fitting factor rankings would be independent of the sequence of the data. Note that several of the clusters



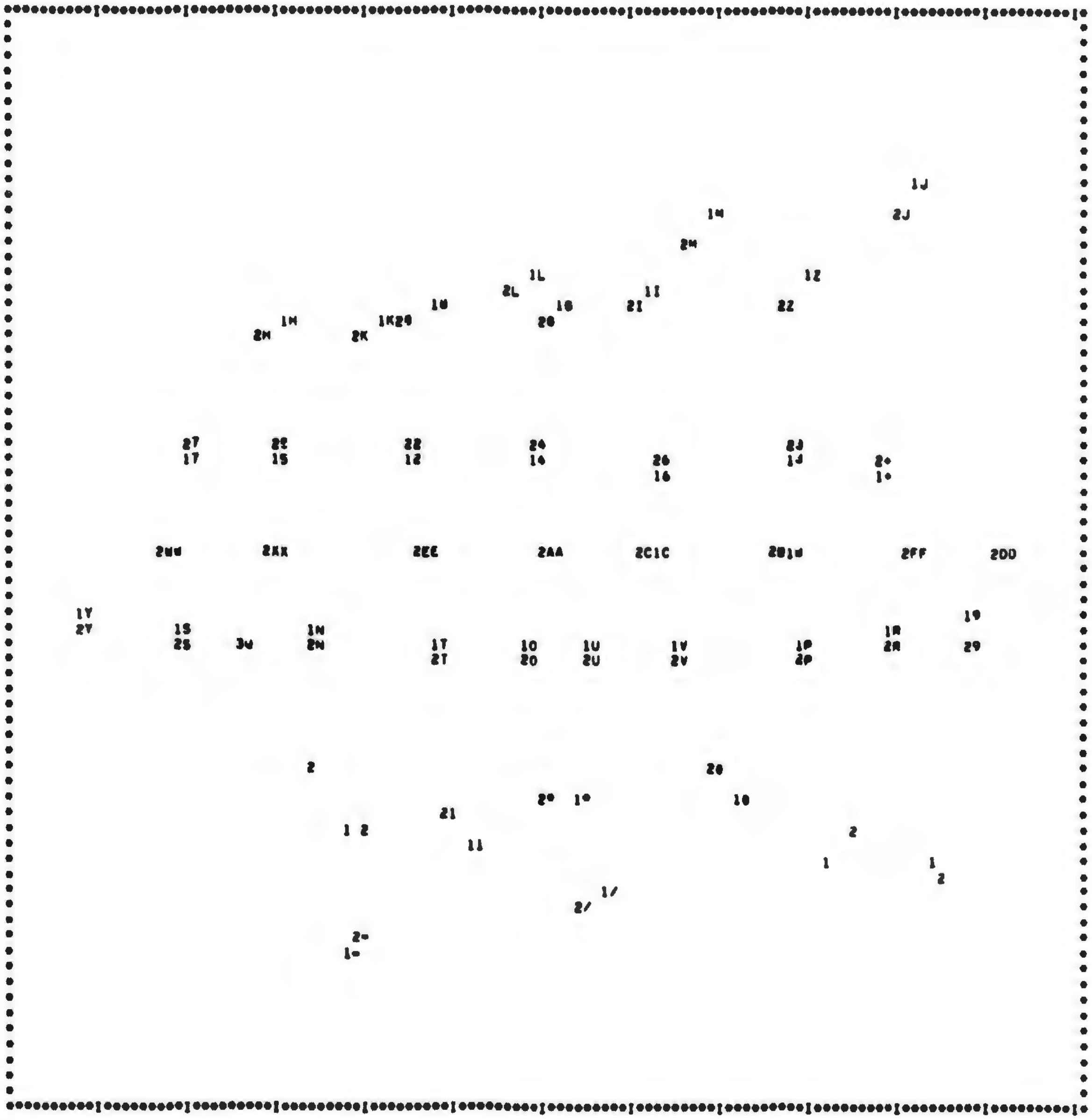
```
CONVERGED AFTER  3 ITERATIONS
```

FIGURE 9 PRINT PLOT WITH ISODATA CLUSTER CENTERS SUPERIMPOSED



PLOT OF VECTORS RESULTING FROM PAIRING OF CLUSTERS WITH THE SAME GENERATED CLUSTER NUMBER.

FIGURE 10 VECTORS RESULTING FROM PAIRING OF CLUSTERS WITH THE SAME CLUSTER NUMBER



```
CRITICAL VALUE =   11.5  NUMBER OF PAIRS MATCHED =  40
```

FIGURE 11   MOTION VECTORS AFTER FOUR ITERATIONS OF THE MOTION PROGRAM, PICTURES 1 AND 2 in the lower left corner of the figure have not been paired because the fitting factor for these pairs has dropped below a predetermined and somewhat arbitrary threshold of confidence.

Figure 12 shows a similar result for Pictures 2 and 3, whereas Figure 11 was for Pictures 1 and 2. In this case there is less confusion of the pairing in the lower left of the picture. The motion vectors found objectively are in general agreement with the motions created artificially, but a more crucial evaluation is given by the SRI cloud console comparison.

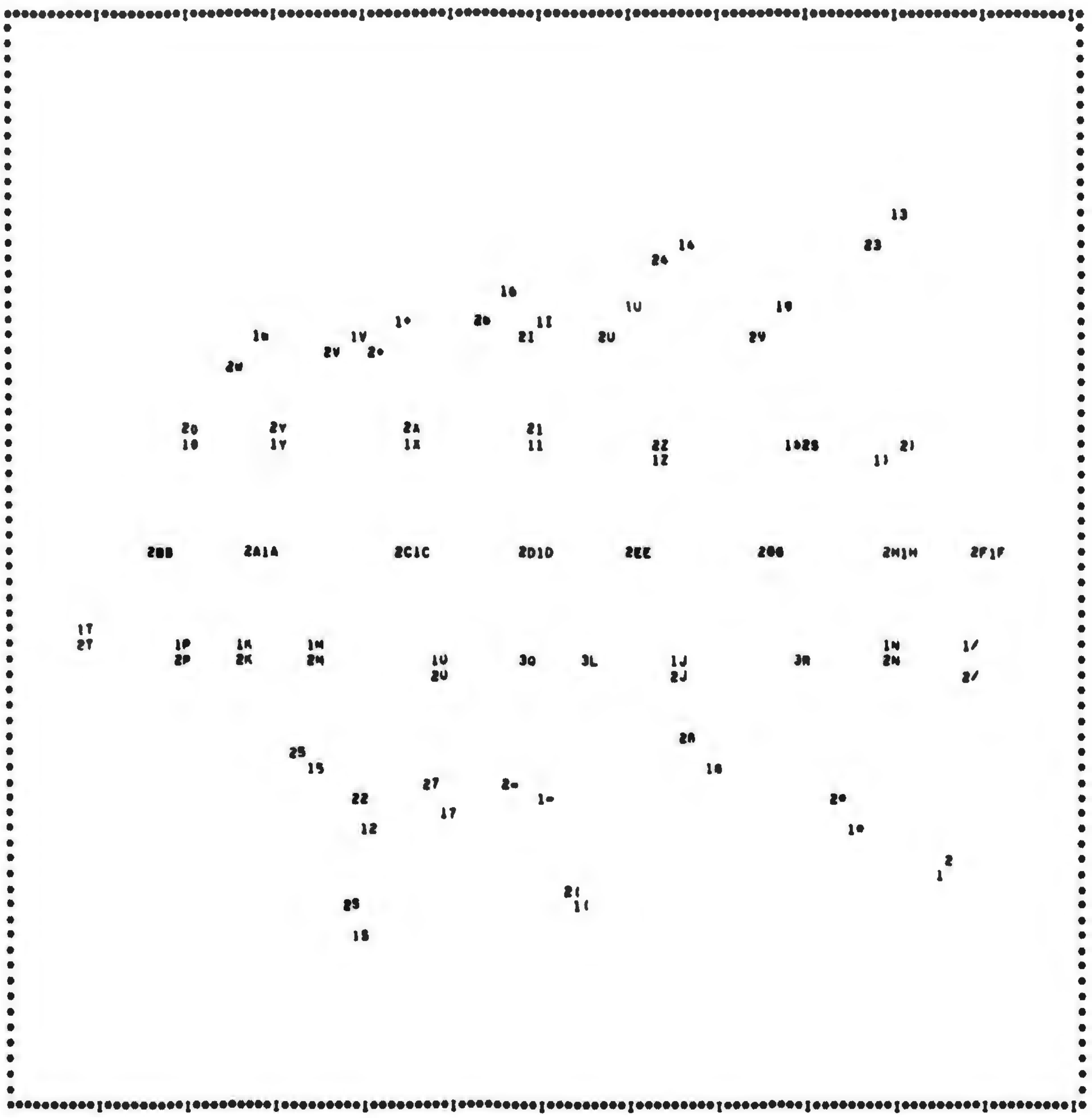
C. Verification of Artificial Motions by SRI Cloud Console

The artificial cloud motions were displayed on the PDP 11 display console and photographed as the example illustrated in Figure 8. The three successive pictures were then read into the memory disc of the SRI cloud console for time-lapse presentation. When viewing these pictures, which were presented in rapid succession, a definite impression of several different motions is obtained which cannot be gained from any amount of static viewing of the pictures.

We did not want to influence the cloud-tracking operators, Messrs. Hadfield and Wiegman, since we wanted them to use their usual approach to cloud tracking. However, before applying their usual approach, they immediately made the following observations on several features of the artificial data:

(1) The clouds move in several basic motions, so they could characterize these by about five basic vectors, although they could easily identify a much larger number of clusters.

(2) The clouds do not change shape from one picture to the next, as do real clouds.



```
CRITICAL VALUE =   11.5  NUMBER OF PAIRS MATCHED =  43
```

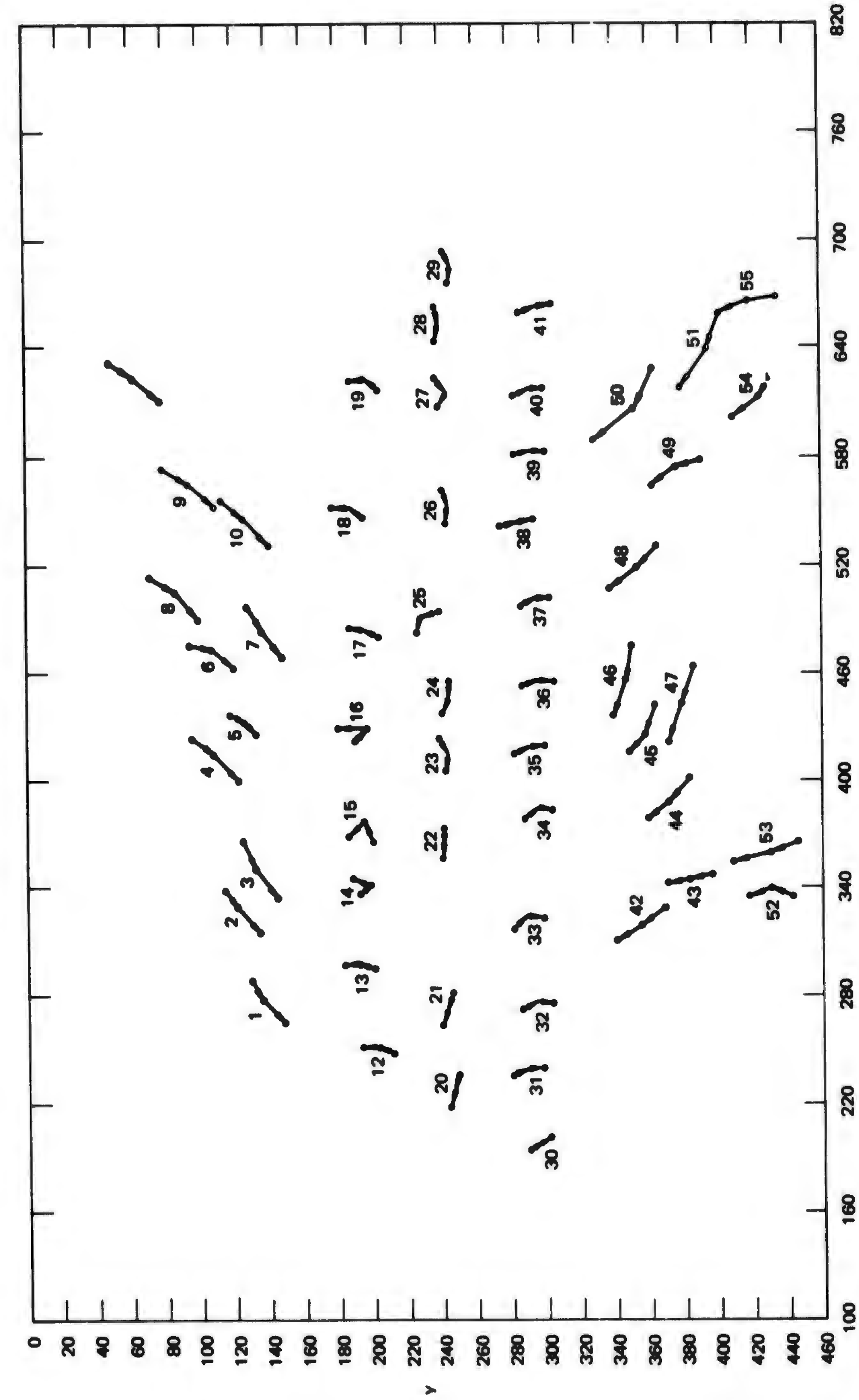
FIGURE 12 MOTION VECTORS AFTER FOUR ITERATIONS OF THE MOTION PROGRAM, PICTURES 2 AND 3

(3) They could estimate cloud centers and track these, or they could locate specific points in the clouds and track them because they can be more accurately identified since they did not appear to change their configuration with respect to an imagined cloud center.

It was agreed to track according to the latter criterion (points) and interpret the results as a cross-check on the accuracy of:

- PDP 11 display plotting.
- Photographic processes.
- Registration processes.
- Human tracking on the SRI cloud console (since the specific points could be accurately specified from the computer commands).

The results of the human cloud tracking are shown in Figure 13; 55 specific data points were tracked. The scaling on this figure is in cloud-console coordinates, whereas the display coordinates of the PDP 11 were used for Figure 4 (with which comparison should be made).



SA-1006-18

FIGURE 13 HUMAN TRACKING OF ARTIFICIAL CLOUDS ON THE SRI CLOUD CONSOLE

VI CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

Our overall conclusion from this work is that it is useful and instructive to study artificially generated cloud data in the manner we have contrived in this project. This is because such data present the essence of the cloud motion analysis problem in a simple and controlled form. As such, we are able to test and improve on as yet relatively unsophisticated techniques in an efficient manner without the distractions and uncertainties presented by real data.

Our interpretation of this comparison is that the objective method using the collection of ISODATA programs gives a satisfactory result compared to human tracking on this set of data. We have not included some of the more realistic and complex features of real data in these examples, but we conclude that our objective methods can extract many significant cloud vectors automatically from such artificial cloud pictures. The same programs used here were used to process real satellite picture data on another project.[5]

In future work, we will generate more complex motions to simulate clouds more realistically. For example, cloud shape changes, and the appearance and disappearance of clouds, are quite likely to happen during the 25 minutes that elapse between successive pictures. The current data-generating language is quite capable of expressing these complexities, although each detail adding to the complexity requires an extra statement in the language. There seems to be no significant way to improve the language without losing the ability to specify such detail. As this research becomes more advanced there may be a need to design a language that can specify large wind flow-fields of various meteorological types.

Preceding page blank

The programs should be enhanced to allow generation of a statistical spread of values for brightness and IR, rather than the single, constant values given for each cluster, as currently implemented. This would be a more realistic situation, and would require specification of standard deviation values for these quantities, as is now done for cluster center position.

A program should be written to automatically plot the motion vectors from the data-generating language sets for several pictures. This corresponds to an automatic version of Figure 4, and also relates to the suggestion above for the specification of wind flow-fields.

A better way to display the data should be found. The current use of print-plots will be inadequate when we add IR data and use multi-layered motions. The PDP 11 display is of good quality and interactive, but our computations are done on the CDC 6400, and that computer has no interactive display. A novel way of using print-plots may have to be devised.

The multiple execution capability of MOTION should be refined to allow a splitting off of different motion vectors in multilevel cases. The IR values should aid in specifying the various levels.

More detailed evaluation should be made of the dot product and other elements of the fitting factor. Also, the dot product is now bivariate, and brightness and IR should be added.

REFERENCES

1. R. M. Endlich, et al., "Objective Methods for Registering Landmarks and Determining Cloud Motions from Satellite Data," Final Report, U.S. Navy Contract N62306-71-C-0068, 29, Stanford Research Institute, Menlo Park, California (October 1971).

2. D. E. Wolf, R. M. Endlich, and D. J. Hall, "Further Development of Objective Methods for Registering Landmarks and Determining Cloud Motions from Satellite Data," Scientific Report, U.S. Navy Contract N62306-71-C-0068, p. 68, Stanford Research Institute, Menlo Park, California (September 1972).

3. R. M. Endlich, et al., "A Pattern Recognition Technique for Determining Cloud Motions from Sequences of Satellite Photographs," Final Report, U.S. Navy Contract N62306-69-C-0312, Stanford Research Institute, Menlo Park, California (April 1970).

4. D. J. Hall, et al., "Development of New Pattern Recognition Methods," 1973. Final Report, SRI Project 1340, Wright Patterson Air Force Base, Ohio, Contract F33615-71-C-1894.

5. D. E. Wolf, et al., "Development of an Automatic Computer System for Measuring Cloud Motion Vectors," Final Report, NASA Contract NAS5-21776, Stanford Research Institute, Menlo Park, California (September 1973).

6. G. H. Ball and D. J. Hall, "ISODATA--A Self-Organizing Computer Program for the Design of Pattern Recognition Preprocessing," Proc. IFIP Congress 65, Vol. 2, pp. 329-330 (May 1965).

ACKNOWLEDGMENTS

We wish to acknowledge the efforts of Roy Endlich in initiating this work and carrying out some of the early research. Unfortunately, he has been away from work with an illness, but we have benefited greatly from informal discussions with him regarding many aspects of our research.

We also acknowledge the efforts of Rex Hadfield and Eldon Wiegman in tracking the artificial cloud data on the SRI cloud console.

Preceding page blank