DTIC FILE COPY

(2)

AFHRL-TR-88-16

AD-A199 403

AIR FORCE HUMAN RESOURCES LABORATORY

MODELING EYE MOVEMENT SEQUENCES USING CONCEPTUAL CLUSTERING TECHNIQUES

Michael S. Belofsky
Don R. Lyon

University of Dayton Research Institute
300 College Park Avenue
Dayton, Ohio 45469

OPERATIONS TRAINING DIVISION
Williams Air Force Base, Arizona 85240-6457

August 1988

Final Report for Period October 1985 - December 1987

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

88 8 2

PAUL CHOUDEK, 1Lt, USAF
Contract Monitor

DEE ANDREWS, Technical Advisor
Operations Training Division

HAROLD G. JENSEN, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S): AFHRL-TR-88-16
6a. NAME OF PERFORMING ORGANIZATION: University of Dayton Research Institute; 6b OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Operations Training Division
6c. ADDRESS (City, State, and ZIP Code): 300 College Park Avenue, Dayton, Ohio 45469	7b ADDRESS (City, State, and ZIP Code): Air Force Human Resources Laboratory, Williams Air Force Base, Arizona 85240-6457
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Air Force Human Resources Laboratory; 8b OFFICE SYMBOL (If applicable): HQ AFHRL	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F33615-84-C-0066, F33615-87-C-0012
8c. ADDRESS (City, State, and ZIP Code): Brooks Air Force Base, Texas 78235-5601	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 62205F; PROJECT NO 1123; TASK NO 03; WORK UNIT ACCESSION NO 83

11. TITLE (Include Security Classification)
Modeling Eye Movement Sequences Using Conceptual Clustering Techniques

12. PERSONAL AUTHOR(S)
Belofsky, M.S., & Lyon, D.R.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Oct 85 TO Dec 87	August 1988	16

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
12	09	
06	04	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
artificial intelligence, cognitive modeling, rule-based systems, concept learning, eye movements, segmentation, conceptual clustering, grouping, visual monitoring.

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

An algorithm for clustering noisy continuous numeric data was developed in a learning system called 2DCG (two-dimensional cluster generalization). The 2DCG system operates in a two-dimensional space, but a general system could operate in an N-dimensional space. The objective of the system was to learn a set of rules which modeled human observers (in the application presented here, this model predicted changes in the eye position of human observers during a visual monitoring task). The rule set had to be complete, consistent, and nonredundant, while minimizing the number and maximizing the generality of the rules. The development of this model and its performance in accounting for noisy data are described.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL: Nancy J. Allin, Chief, STINFO Office	22b TELEPHONE (Include Area Code): (512) 536-3877; 22c OFFICE SYMBOL: AFHRL/TSR

DD Form 1473, JUN 86 — Previous editions are obsolete — SECURITY CLASSIFICATION OF THIS PAGE

Unclassified

SUMMARY

The goal of this 6.1 research was to develop methods for predicting the location of cognitively driven eye movements during a difficult visual monitoring task. If eye movements can be accurately predicted in such situations, then methods might eventually be developed for diagnosing and training optimal allocation of attention during visual monitoring tasks. This report is a brief summary of the development and evaluation of one promising prediction method, two-dimensional cluster generalization (2DCG). The data to be predicted were changes in the eye position of observers viewing a display of four rapidly changing numbers representing aircraft instrument values during level flight. The monitoring task was to detect the first number to fall outside a window of tolerable values. Each change in eye position from one number to another was recorded, together with the most recent value of the observed number, the amount by which that value had changed during observation, and the time (in milliseconds) that the number had been observed. These variables formed the axes of a set of two-dimensional spaces into which the data were plotted. An algorithm was developed for partitioning these spaces into regions that predicted a change in eye position to a particular next "instrument." The resulting model correctly predicted about 75% of the changes in eye position in the data on which it was developed and about 50% of eye position changes in new cross-validation data.

PREFACE

This 6.1 research was performed at the Air Force Human Resources Laboratory (AFHRL) at Williams Air Force Base and was supported by the Air Force Office of Scientific Research (Life Sciences Task 2313T3) and the Air Force Human Resources Laboratory (Contract F-33615-84-C-0066 to the University of Dayton Research Institute). We would like to thank David Benjamin (University of California at Irvine) and Cindy Arroyo (UDRI) for providing suggestions on the content of this paper, and Dr. Elizabeth Martin (AFHRL) for helpful comments at various stages of the research.

TABLE OF CONTENTS

I. INTRODUCTION

Conceptual clustering systems, such as CLUSTER/2 (Michalski & Stepp, 1983), RUMMAGE (Fisher, 1984), and COBWEB (Fisher, 1987), group together similar objects, each described by values for a fixed set of attributes, (e.g., for soybean classification: precipitation = low, root-condition = rotted). The degree of similarity of these objects is determined by an evaluation function which examines appropriate discrete attribute values. Conceptual clustering systems effectively categorize the input objects into distinct classes based on their attribute values. Input objects with similar attribute values are grouped into the same class.

The systems cited above all work with attributes having a small number of discrete values. When attributes have continuous (or real) values, clustering becomes difficult because two values which are numerically close may be placed in two different clusters. Clustering of continuous data is possible by creating discrete categories from ranges of numeric values based on the best distribution of the data. Lebowitz's (1985) UNIMEM system handles such continuous numeric data.

In our project, a conceptual clustering system that operates on continuous numeric attributes was implemented. The system's name, 2DCG, stands for "two-dimensional cluster generalization." 2DCG operates in a two-dimensional space, creating clusters from discrete ranges of values which are determined by the data rather than given to the system. The input data are generalized in the sense that two or more non-adjacent real values can form a cluster if there are no intervening clusters. 2DCG is a non-incremental learning system that operates on data about the locations of eye movements made by human observers trying to monitor four simultaneously displayed numbers with rapidly changing values. The system must extract a set of rules (described below) which predict these changes in eye position. Each change in eye position is predicted from the information available on the instrument on which the observer is currently fixated. (Several attempts to use the previous instruments that the observer had fixated were unsuccessful because the resulting rules were too specific.)

A rule tells the system what to do next (in this system, which instrument to fixate next) based on the current data available (the information contained at the currently viewed instrument location). These rules are in the standard IF-THEN format. The IF part of a rule contains tests for the instrument that is being fixated. Along with the instrument's name, the IF part contains the value of the instrument when it is first fixated, the change in the instrument value while it is fixated, and the duration for which the instrument is fixated. The THEN part of a rule contains the name of the instrument that will be fixated next if the information in the IF part is true.

The first value displayed by an instrument that is being fixated and the size of the change in this instrument's value for the duration of the glance (if the change is not zero) form the axes of a two-dimensional space of predictive conditions. The conceptual clustering techniques, which are described later, partition this space into the smallest possible clusters with the most evidence to support each cluster. There are no overlapping rules; if two (or more) clusters overlap, the cluster with the most evidence remains as a cluster and eventually becomes a rule, and the other clusters are deleted. This process guarantees that the rule base is consistent. It is also nonredundant because only one rule remains at any single position in the two-dimensional space. Finally, similar adjacent clusters are combined, and then built up or "grown" so that all possible combinations of attribute values are accounted for. This step guarantees that the rule base is complete. These rules are presented to a production system which executes the appropriate rule given each scenario. Section II discusses some of the background that motivated this research. Section III describes the conceptual clustering techniques involved in the 2DCG system. Section IV presents the algorithm used in the 2DCG system. Section V examines the performance of this system using three human observers. Finally, Section VI provides some conclusions.

II. BACKGROUND ON THE VISUAL MONITORING TASK

The goal of this research is to develop and test a model which predicts changes in the eye positions of human observers monitoring multiple sources of information. Each change in eye position from one information source to another is called a glance transition. The information sources are numbers representing the values of four aircraft instruments. These numbers are driven by flight equations from a T-38 aircraft and are presented in the four corners of a display controlled by an IBM XT. The observer's task is to monitor the four "instruments" and respond when any one of them exceeds particular critical values.

The current research deviates significantly from past models of eye fixations in monitoring tasks, such as models by Ellis and Stark (1981) and Senders (1983). The model presented here predicts individual glance transitions between instruments, whereas most previous models have predicted the aggregate frequency with which an instrument would be monitored or the average duration of a glance at a particular instrument. All of these models were probabilistic, whereas the system presented here creates a deterministic model.

III. DEVELOPMENT OF THE 2DCG SYSTEM

In order to create a computer model of glance transition patterns, several tasks were performed. First, a computer simulation of a T-38 aircraft, which presented the instrument values to the human observers, was implemented. Second, eye position data were collected and initial data smoothing was performed to reduce noise. Finally, rule generation and generalization algorithms were implemented to analyze the data collected from the experiment above. These algorithms form the 2DCG system. A more detailed description of these tasks is presented in Belofsky (1987) or Belofsky and Lyon (1988).

The data collected from this experiment are (a) the eye positions of the observer at discrete sample times and (b) the corresponding instrument values. Each adjacent pair of data points is then converted to glance transitions, which are the input data to the 2DCG system. The 2DCG system utilizes one-quarter of the transition data for each observer to create a rule-based model of eye movement sequences. This model is cross-validated on the other three-quarters of the transition data.

Two attributes of the instrument data are used in the clustering algorithm of the 2DCG system. The first attribute is always the instrument's value when the observer initiates the glance at that instrument. The second attribute is sometimes the amount of change in the instrument's value, the delta value, for the duration of the glance. However, if the instrument's value is stationary for the duration of the glance, a third attribute, the glance duration, is used rather than the delta value. This third attribute was found empirically to provide more information when the delta value was zero.

Conceptually, either of these pairs of attributes forms a two-dimensional space in which the continuous numeric data are distributed. Forming clusters in this space is a problem similar to image processing problems, in which pixels with similar values are combined to form larger regions. In this analogy, the regions (clusters) into which the pairs of attributes (pixels) are divided form different "pixel values." These regions eventually become rules in the two-dimensional space.

The system presented here operates in a two-dimensional space because information required to model the observers effectively could be represented in this way. There is no reason why this system cannot be extended to operate in a three-dimensional space or more. The techniques developed for this system are capable of operating in an N-dimensional space; however, complexity will naturally increase with larger spaces.

A typical transition entered into the 2DCG system can be paraphrased as:

> If indicated airspeed is 701 and changed by 1 during the 625 milliseconds it was fixated, then pitch was looked at next.

Table 1 presents some sample data for the indicated airspeed instrument, with delta values. For this system, there are eight two-dimensional spaces that are addressed separately: each instrument with delta values and each instrument with time values (when the delta values are zero). Indicated airspeed was chosen to illustrate the clustering techniques because it encompasses a smaller range of values.

In earlier versions of this system, the clusters found were the ones that covered the largest possible contiguous area and encompassed the most transitions. This caused a major problem. In some cases, clusters grew to cover all instrument and/or delta (time) values. This created a model which was too general in its predictions and thereby misclassified the data. Smaller clusters within the larger clusters, which were often better predictors in their localized area, were not identified. That is, local maxima were ignored in favor of larger clusters. Therefore, an algorithm was developed to find small clusters and gradually build them up to cover the entire data space.

In this system, all clusters initially consist of one instrument value and one delta (or time) value. At each instrument value and delta (or time) value, the instrument that is transitioned to most often is selected as the output of that cluster. This output is assigned a weight equal to the number of transitions that occur to the selected instrument. From any instrument value and delta (or time) value, a transition can occur to one of the three other instruments. If several instrument transitions have the same maximum weight, one is arbitrarily chosen. Therefore, at each position in the two-dimensional space, the data presented in Table 1 are converted to the information in Table 2. Table 2 contains the smallest clusters at each coordinate where transitions occur in the data. Each cluster consists of one transition with its corresponding weight.

This process guarantees that the model created by the 2DCG system predicts the maximum number of transitions on which this model is created. This model does not require conflict resolution techniques while it is executing because it is created such that no conflicts ever occur. The model is, therefore, deterministic; whenever a specific instrument value and delta (or time) value occur, the same transition is predicted.

The clusters as presented in Table 2 show that there are many empty regions which have to be filled in, in order to guarantee complete coverage. The empty regions are filled in by comparing the weights of all clusters in the immediate vicinity. The weights of the clusters immediately surrounding an empty region are totaled. The instrument that is transitioned to most is placed in the empty region. For example, in Table 2 at instrument value 701 and delta value -2, the pitch (P) instrument will be transitioned to next because the weight for the transitions to the pitch instrument in clusters immediately surrounding the above empty region totals to 16, while the weight for the altitude (A) instrument totals only to 6. Adjacent clusters which have the same transition instrument are then combined together to form larger clusters. Table 3 shows what the data look like for the indicated airspeed instrument using delta values after this step is completed.

The clustering procedure described thus far is very sensitive to (random) noise in the data. All transitions from the input data are used to create the model; therefore, clusters with low weights become rules when they may really be noise. The data are collected from human observers (subjects) and contain noise because the observer either blinks or moves out of the recording area momentarily. In order to reduce the amount of noise incorporated into the rule base, all clusters with weights less than 5 are removed before any processing is done to the clusters. This minimum number was found empirically by creating several different models with different minimum-sized clusters and examining the prediction percentages. The prediction rate reached asymptote when the minimum size of a cluster was 5.

The noise that is in the system before the above technique is employed is amplified by the cluster-growing techniques because if a cluster is really noise and is surrounded by empty regions, these empty regions will be filled by this erroneous cluster. This lowers the prediction rates on the cross-validation data. There is still noise in the data, but much of the noise has been removed by filtering out clusters with lower supporting weights.

IV. THE ALGORITHM

The previous section presented a description of how the 2DCG system creates rules which model the observers for the aircraft instrument domain. Figure 1 presents the general algorithm to cluster data in an N-dimensional space.

---

* Distribute the data that will be used to cluster, into the appropriate positions in the N-dimensional space.
* At each position in the N-dimensional space:
    * If there is more than one action to take, choose the one with the most weight.
    * If two actions have the same weight, arbitrarily choose one.
* If there is some minimum weight, K, that each individual position in the N-dimensional space must be greater than, remove each data point whose weight is less than K.
* Until there are no empty positions in the N-dimensional space:
    * Total weights for each different action for the immediately surrounding positions (two in a one-dimensional space, eight in a two-dimensional space, 26 in a three-dimensional space, etc.)
    * The adjacent position that has the most weight absorbs the empty position. If two or more actions have the same weight, arbitrarily choose one. This newly absorbed region's weight will be zero.
    * If there are no non-empty positions immediately surrounding the empty region in question, compare the number of newly absorbed positions for each action immediately surrounding the empty position. The action that has the most supporting positions is placed into the empty region.
* Combine adjacent positions which have the same action to form larger clusters.
* Convert above clusters into rules based on the area in the N-dimensional space they occupy.

Figure 1. General Algorithm to Cluster Data in an N-Dimensional Space.

V. THE PERFORMANCE OF THE 2DCG SYSTEM

Three observers were modeled to test the performance of the 2DCG system. One was a T-38 pilot; the other two were non-pilots. The pilot was expected to be much more consistent than the non-pilots and he was, but not as much as expected. On the average, the model is capable of predicting 65% of the dominant transitions. The dominant transition is the instrument which is transitioned to most frequently in a given situation. That is, at any one position in the two-dimensional space, there can be a transition to one, two, or all three of the other instruments. For example, for the indicated airspeed instrument at instrument value 699 and delta value -2, if there are five transitions to pitch (P), four transitions to vertical velocity (V), and eight transitions to altitude (A), the dominant transition is from the indicated airspeed instrument to the altitude instrument. The weight of this dominant transition is eight.

The model created by the 2DCG system is, by definition, capable of predicting 100% of the dominant transitions on the data used to create it, when clusters of all sizes are incorporated into the model. When the smaller clusters are filtered out, the prediction rate falls below 100% for the creation data set because the model is no longer recognizing regions of lower weight.

Table 4 presents the prediction percentages for the three observers using rules created with clusters of any size. For non-pilot observer SN, there are 949 transitions in the data used to create the model (the DATA data set). The model is created using these data, and 733 of these transitions are predicted, which is 77% of the data. Because the model is created on these data, 733 is also the maximum number of dominant transitions; therefore, the model predicts 100% of the dominant transitions.

The transition set labeled "ODST" is one of the three cross-validation data sets. Using the model created on the DATA data set, 396 transitions are predicted. This accounts for 51% of the total transitions. In the ODST data set, 622 transitions (80%) are dominant. Therefore, the model predicts 64% (396/622) of the dominant transitions for this data set.

Table 4 also presents the prediction percentages using the 2DCG system with rules created for clusters having a minimum weight of 5. Notice that while the prediction percentages drop a little for the data on which the model is created, the cross-validation prediction percentages increase. This suggests that low-weighted clusters are likely to be noise. By removing these low-weight clusters, the prediction percentage falls on the data used to create the model, but this indicates that the model is becoming less specific to these data and more capable of predicting transitions that it has never seen before.

VI. CONCLUSION

A conceptual clustering system which learns to model human eye positions in a simulated aircraft instrument monitoring task was presented. This system, 2DCG, operated on continuous numeric attributes with random noise. The attribute values were clustered in a two-dimensional space, although the system could be extended to work in an N-dimensional space.

Evaluating the performance of the model created by the 2DCG system in this context is difficult, because the number of glance transitions predicted is dependent upon the consistency of the data that are being modeled. However, it may be a promising candidate for evaluation using artificial continuous numeric data generated using predefined clusters plus varying amounts of random noise. Advances in the technology of predicting glance transitions will require learning techniques that are very robust when given noisy data.

Table 1. Transition Data for the Indicated Airspeed Instrument

Graphical Representation:

Instrument Value

Delta Value	695	696	697	698	699	700	701	702	703	704	705
-2	A 4			A 7 P 4	A 8 P 5 V 4	A 6 P 4		A 4 P 6	P 5		V 4
-1	A 5 P 4	A 4		A 5	A 6 P 5	A 4 P 6		P 4 V 3	A 5 P 5		V 5
1	A 5	A 6	A 4 P 4 V 4	A 5	A 5 P 4	A 4 P 3	A 5 P 5	P 3	A 6 V 7	V 4	V 5
2	A 2	A 4			A 4	P 5	P 6	P 5		V 5	V 3

Interpretation:
When the indicated airspeed's instrument value is 699 knots and its delta value is -2, there are eight transitions to altitude, five transitions to pitch, and four transitions to vertical velocity.

Table 2. Clusters for the Indicated Airspeed Instrument

Instrument Value

Delta Value	695	696	697	698	699	700	701	702	703	704	705
-2	A 4			A 7	A 8	A 6		P 6	P 5		V 4
-1	A 5	A 4		A 5	A 6	P 6		P 4	P 5		V 5
1	A 5	A 6	A 4	A 5	A 5	A 4	A 5	P 3	V 7	V 4	V 5
2	A 2	A 4			A 4	P 5	P 6	P 5		V 5	V 3

Table 3. Final Clusters for the Indicated Airspeed Instrument

Instrument Value

Delta Value	695	696	697	698	699	700	701	702	703	704	705
-2	A	A	A	A	A	A	P	P	P	P	V
-1	A	A	A	A	A	P	P	P	P	V	V
1	A	A	A	A	A	A	A	P	V	V	V
2	A	A	A	A	A	P	P	P	V	V	V

Table 4. Two-Dimensional Cluster Generalization Model Prediction Percentages

	DATA	ODST	SDOT	ODOT
Non-Pilot Observer SN:				
TRANSITIONS	949	776	826	711
PREDICTED	733 77% *682 72%	396 51% *411 53%	403 49% *421 51%	352 50% *362 51%
DOMINANT	733 77%	622 80%	650 79%	571 80%
DOMINANT TRANSITIONS PREDICTED	733/733 100% *682/733 * 93%	396/622 64% *411/622 * 66%	403/650 62% *421/650 * 65%	352/571 62% *362/571 * 63%
Non-Pilot Observer TN:				
TRANSITIONS	1073	1080	1008	1134
PREDICTED	811 76% *747 70%	493 46% *509 47%	450 45% *467 46%	505 45% *512 45%
DOMINANT	811 76%	826 76%	783 78%	856 75%
DOMINANT TRANSITIONS PREDICTED	811/811 100% *747/811 * 92%	493/826 60% *509/826 * 62%	450/783 57% *467/783 * 60%	505/856 60% *512/856 * 60%
Pilot Observer PL:				
TRANSITIONS	1104	1134	1221	1261
PREDICTED	856 78% *799 72%	584 52% *619 55%	589 48% *617 51%	665 53% *709 56%
DOMINANT	856 78%	858 76%	894 73%	945 75%
DOMINANT TRANSITIONS PREDICTED	856/856 100% *799/856 * 93%	584/858 68% *619/858 * 72%	589/894 66% *617/894 * 69%	665/945 70% *709/945 * 75%

* Minimum Cluster Size = 5

REFERENCES

Belofsky, M. (1987). Artificial-intelligence-based models of eye movement sequences during a visual monitoring task. Master's Thesis, Computer Science Department, Arizona State University.

Belofsky, M., & Lyon, D. R. (1988). Artificial-intelligence-based models of eye movement sequences during a visual monitoring task (UDR-TR-88-29). Dayton, OH: University of Dayton Research Institute.

Ellis, S. R., & Stark, L. (1981). Pilot scanning patterns while viewing cockpit displays of traffic information. Proceedings of the 17th Annual Conference on Manual Control, 517-524.

Fisher, D. (1984). A hierarchical conceptual clustering algorithm (Technical Report). University of California, Irvine: Department of Information and Computer Science.

Fisher, D. (1987). Conceptual clustering, learning from examples, and inference. Proceedings of the Fourth International Workshop on Machine Learning, 38-49.

Lebowitz, M. (1985). Categorizing numeric information for generalization. Cognitive Science, 9, 285-308.

Michalski, R., & Stepp, R. (1983). Automated construction of classifications: Conceptual clustering versus numerical taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No. 4, 396-410.

Senders, J. W. (1983). Visual scanning processes. Drukkerij: Neoprint, Soest.

* U. S. GOVERNMENT PRINTING OFFICE: 1988