

DAA/CODDARD

NAS5-29269

# PATTERN RECOGNITION OF SATELLITE CLOUD IMAGERY FOR IMPROVED WEATHER PREDICTION

IN-47

64419-CR

P.18

## FINAL REPORT TO NATIONAL AERONAUTICS AND SPACE ADMINISTRATION PHASE I SBIR CONTRACT No. NAS5-29269

**CHASE CONSULTING, INC.  
3543 Caminito Carmel Landing  
San Diego, California 92130**

by

**Catherine Gautier  
Earth Space Research, Inc.**

**Richard C.J. Somerville  
Oomaw Corporation**

**Leonid B. Volfson  
Chase Consulting, Inc.**

**September 1986**

(NASA-CR-180720) PATTERN RECOGNITION OF  
SATELLITE CLOUD IMAGERY FOR IMPROVED WEATHER  
PREDICTION Final Report (Chase Consulting)  
18 p Avail: NTIS HC A02/MF A01 CSCL 04B

N87-29987

Unclas  
0064419

G3/47

## ABSTRACT

The major accomplishment of this contract has been the successful development of a method for extracting time derivative information from geostationary meteorological satellite imagery. This research is a proof-of-concept study which demonstrates the feasibility of using pattern recognition techniques and a statistical cloud classification method to estimate time rate of change of large-scale meteorological fields from remote sensing data. The cloud classification methodology is based on Typical Shape Function analysis of parameter sets characterizing the cloud fields.

The three specific technical objectives, all of which were successfully achieved, are as follows:

1. Develop and test a cloud classification technique based on pattern recognition methods, suitable for the analysis of visible and infrared geostationary satellite VISSR imagery.
2. Develop and test a methodology for intercomparing successive images using the cloud classification technique, so as to obtain estimates of the time rate of change of meteorological fields.
3. Implement this technique in a testbed system incorporating an interactive graphics terminal to determine the feasibility of extracting time derivative information suitable for comparison with numerical weather prediction products.

This work should be carried forward in Phase II to produce a fully operational stand-alone work station to enable the benefits of this research to be realized in a real-time numerical weather prediction environment. Forecast verification studies, in conjunction with predictability theories, show that the early hours of numerical weather predictions are typically characterized by a rapid decay of skill. If observational estimates of time derivative, can be compared with numerical computations during this critical period, a potentially valuable new tool will have been made available for improving numerical weather prediction model.

## Introduction

The basic concept incorporated in our Phase I proposal was to increase the value and utility of meteorological satellite data for numerical weather prediction by extracting time derivative information from geostationary imagery. In Phase I, we have undertaken a proof-of-concept study to demonstrate the feasibility of using a statistical cloud classification method and pattern recognition techniques to estimate time rates of change of large-scale meteorological fields from remote sensing data.

In Phase I, we have exploited the fact that VISSR images from GOES are typically available at half-hour intervals, but that conventional meteorological analysis techniques produce only two analyses per day, so that much of the information content of these images is not utilized. We proposed to analyze sequences of such images by pattern recognition methods to produce observational estimates of time derivatives of meteorological parameters during the first few hours of the large-scale forecast period, when the skill of the numerical forecast typically decays most rapidly. Such observational time derivative information has great potential for increasing the forecast skill, both by statistical corrections to dynamical forecasts and by physical model improvements.

Our Phase I project had three specific technical objectives:

1. Develop and test a cloud classification technique based on pattern recognition methods, suitable for the analysis of visible and infrared geostationary satellite VISSR imagery.
2. Develop and test a methodology for intercomparing successive images using the cloud classification technique, so as to obtain estimates of the time rate of change of meteorological fields.
3. Implement this technique in a testbed system incorporating an interactive graphics terminal to determine the feasibility of extracting time derivative information suitable for comparison with numerical weather prediction products.

We have successfully completed all of these objectives.

## Major Results

We have worked on these objectives using a series of GOES VISSR images, but we illustrate our results with a case study which is presented in detail below.

We first selected four infrared GOES-6 images for test purposes. Each image consisted of an array of 500 by 500 pixels. Each pixel nominally covers an area of 4 km by 4 km; the size of the area varies with latitude. The images were from November, 1983 and cover an

area in the eastern North Pacific Ocean off the west coast of the United States extending from  $30^{\circ}$  to  $50^{\circ}$ N. The selected images include a number of interesting meteorological features of synoptic scale. We chose to focus on a mature extratropical cyclone as an example of a baroclinic development which would be an important element of a numerical forecast. A cyclone is typically associated with different air masses separated by a front which can be visualized on satellite images as a curved cloud mass. It is this cloud mass that we will be using to determine the location and displacement of the cyclone from the analysis of satellite images.

In keeping with the procedure outlined in our Phase I proposal, we applied the Typical Shape Function (TSF) analysis to a subset of each original image, comprising approximately one third of the original image domain. The TSF analysis is used to classify the set of parameters describing the cloud fields on the image. We found that for the parameters of greatest interest, which include statistics such as contrast and maximum and minimum brightness, we needed to coarsen the original resolution. We finally settled on a new pixel size of 8 km by 8 km, and we computed the parameters in non-overlapping boxes of 4 by 4 new pixels. The domain which we analyzed on each image consists of an array of 32 by 32 of such boxes. For test purposes, we elected to treat a time sequence of four such arrays, at one hour intervals, e.g., at 0, 1, 2, and 3 hours, to investigate the feasibility of extracting time derivative information (the selected images are presented in Figures 1a and 1b).

Each box is characterized in our procedure by a "texture" defined by the values of seven parameters (here we chose not to include visible information because of its lack of availability at night). These parameters are the mean brightness value, minimum and maximum values, mean of the distribution, contrast, entropy and first moment. The appropriateness of this choice of parameters for describing a cloud field has been previously tested by us. This texture is the basis for the classification of each box; each parameter can be viewed as a different measurement of the cloud horizontal properties similar to multispectral measurements. Thus, a "profile" is generated for each box with 7 different "levels" and these profiles are then classified according to their shape, in this case the relationship between the parameters. Each class represents an idealized type of cloud coverage (e.g., clear, partly cloudy, uniform cloud). We have found that the use of texture as a classification variable facilitates the analysis, primarily because it gives more information than would a threshold technique based on a single binary parameter such as brightness. A first application of the TSF analysis to our series of images provided four statistically significant classes (Figure 2). From this figure it is clear that typically two of these classes occur within the cyclone, and two others occur outside it.

To determine the displacement of the cyclone cloud mass over the sequence of images

we have tested two different approaches. The first one is based on the displacement of the edge of the cyclone cloud mass and the second one is based on the displacement of the entire cyclone cloud mass as a feature of homogeneous characteristics (i.e., belonging to one class only).

For the first approach (i.e., edge displacement), we initially merged the pairs of classes (on the classified image of Figure 2) which most resemble one another to obtain only two classes. It then becomes quite easy to determine the boundary of the cyclone clouds on each classified image as the boundary between these two classes. Although the merging was manually performed, it can easily be automated based on criteria of resemblance between the "profiles" characterizing each class. It is the boundary of the cyclone cloud mass that we want to follow to assess the cyclone's displacement over the time interval between two images. Because of the selected size of the original boxes (32 km by 32 km) the resulting boundary has a very coarse resolution which would not allow an accurate determination of its displacement. The accuracy is limited by the size of the pixel used to compute the displacement and by the period of time over which this displacement is estimated. To enhance the resolution of the boundary we reclassify its surrounding area by applying the TSF analysis to a data set that has a higher resolution (e.g., 8 km by 8 km). To automatically determine the region surrounding the edge of the cyclone cloud mass, we have developed an algorithm which is based on a gradient method with specified width around the region of high gradient. The first width tested contained 2 pixels on each side of the edge; the results of this procedure are presented in Figure 3. The second width contained one pixel on each side, to limit the area to be later re-analyzed; the results are presented in Figure 4. A comparison of the two sets of results suggests that the narrower width provides sufficient information for the classification and reduces the processing time (i.e. TSF analysis computing time) by a factor of four. Once this narrow region has been re-classified, we obtain a precisely located edge which can be followed from one image to the next (Figures 5a and 5b).

For the second approach we apply a pattern recognition technique known as the window matching technique to two successive classified images. A window of a limited size is selected on the first image and a search is made for the corresponding window (i.e. with the same brightness characteristics) on the second image. The second window is selected by the simple expedient of maximizing the correlation of the textures (classes). In order to minimize the time for the search we apply an artificial intelligence technique whereby the search is started from a point on the second image which corresponds to the center of the original window. The search continues outward from this point following a spiral pattern with the rate of change of the correlation values being continuously monitored as the search

extends outward. The search is terminated when a near perfect ( $\geq .97$  correlation value) match has been found. As soon as the rate of change of the correlation values becomes negative the search is terminated. Furthermore, the search is also terminated when the displacement becomes non-physical (this limit is provided a priori from meteorological knowledge). The region of maximum correlation value is thus selected as the matching region and the displacement is computed as the difference in location on the two successive images. Obviously, the accuracy of the time rate of change estimates depends on the alignment of the two images. Here we use the navigation parameters provided by NOAA to accurately register each individual image and verify this registration using landmarks and coastal information. This is achieved in a standard manner whereby the location of some coastal points is computed from the given navigation parameters and a coastline graphic is generated from an accurate data base (e.g., CIA orographic data base). This coastline graphic is then overlaid onto the image and the navigation parameters are recomputed until the coastline graphic exactly matches the coastline detectable on the satellite images (i.e., the coastline characterized by a change in visible brightness or surface temperature on the satellite image). This procedure is performed only when a sufficiently large part of the coastline is detectable on the satellite image. Results obtained by applying this method to three pairs of successive classified images are presented in Figures 6, 7 and 8.

### **Future Work**

The computed displacement vectors appear physically reasonable in both direction and magnitude, but the accuracy of the displacement estimations has not yet been validated. We propose to do that during Phase II.

Several refinements of these two approaches are possible. One which seems especially attractive is to improve the definition of the edge (the frontier or border between the cyclone cloud mass and its environment) by performing the TSF analysis using overlapping boxes in the narrow strip or sub-domain which includes the edge.

These results appear to be extremely promising with respect both to the feasibility of the proposed technique and to its potential value for improving large-scale numerical weather predictions. By extracting time derivative information from sequences of geostationary images, we are recovering valuable observational information which has not previously been available to the analysis-forecast procedure. We conclude, on the basis of the Phase I results, that the original concept is indeed viable, and that proceeding to develop it to a practical technology in Phase II is justified. The Phase I results have also pointed the way to a number of improvements and extensions which we plan to carry out

during Phase II.

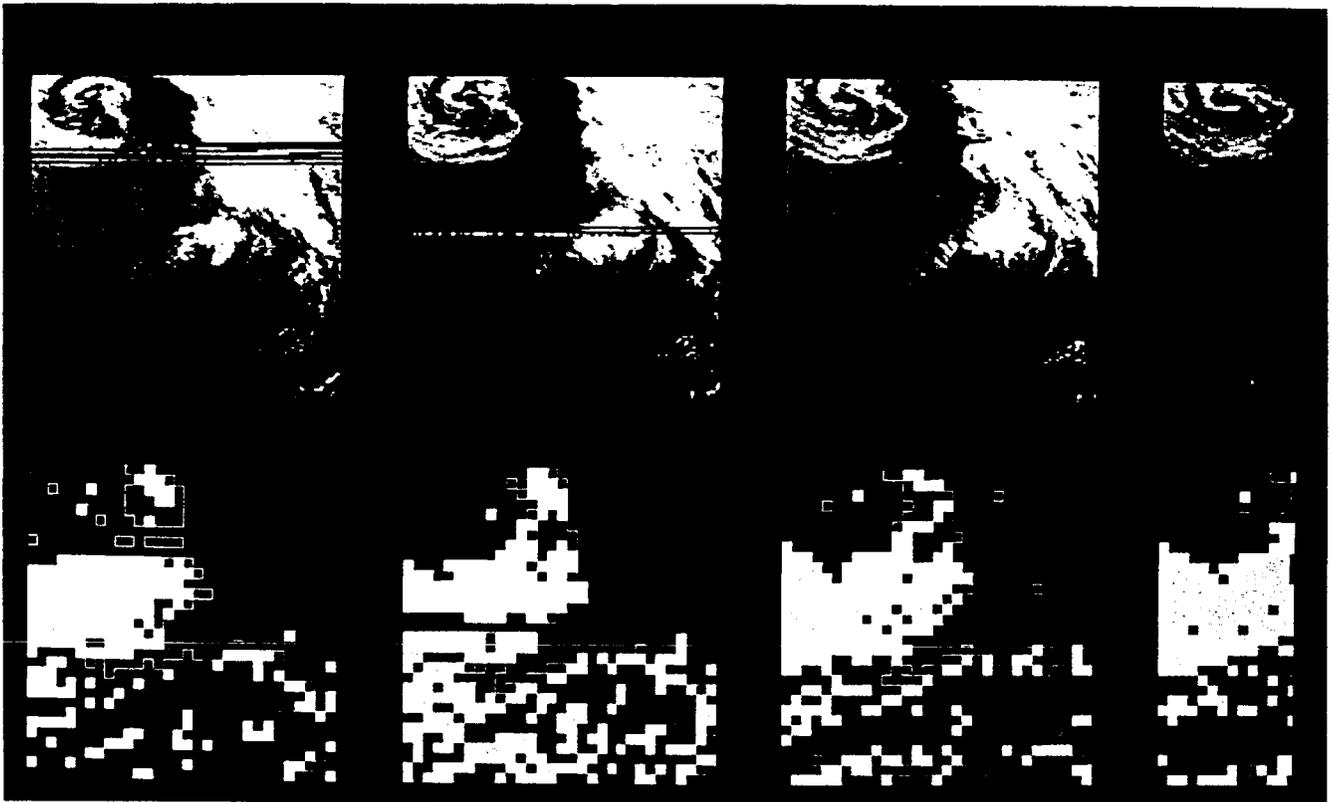


Figure 1a



Figure 1b



Figure 2

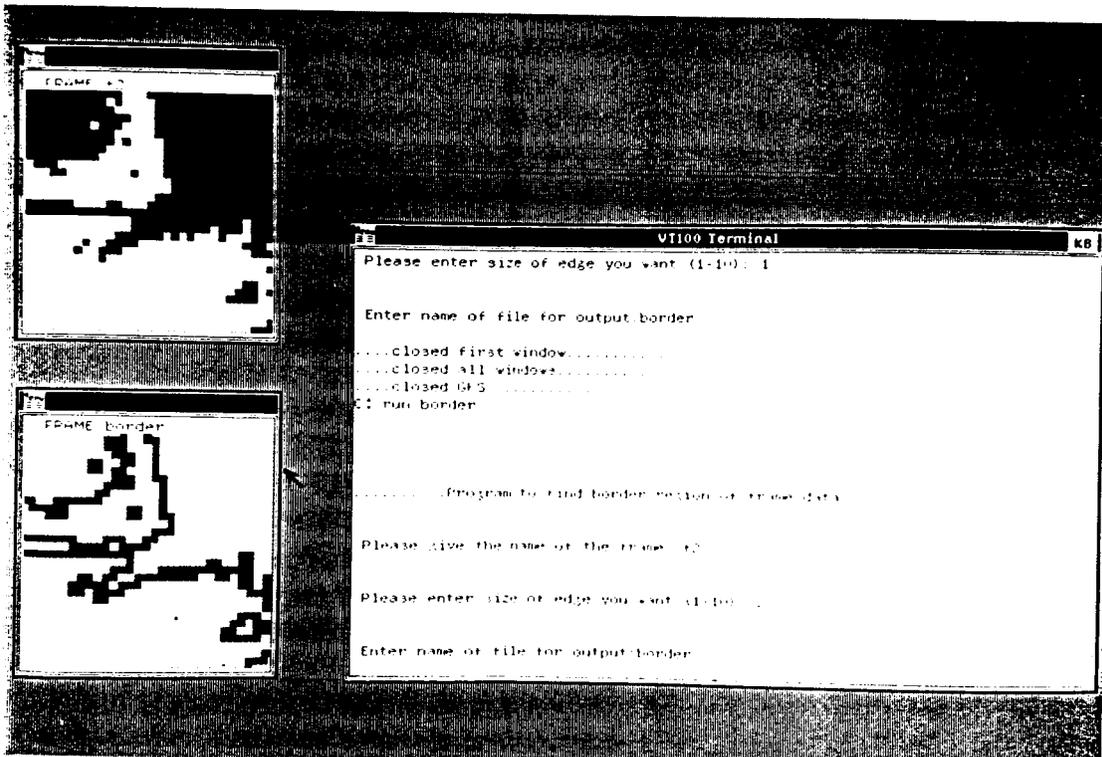


Figure 3

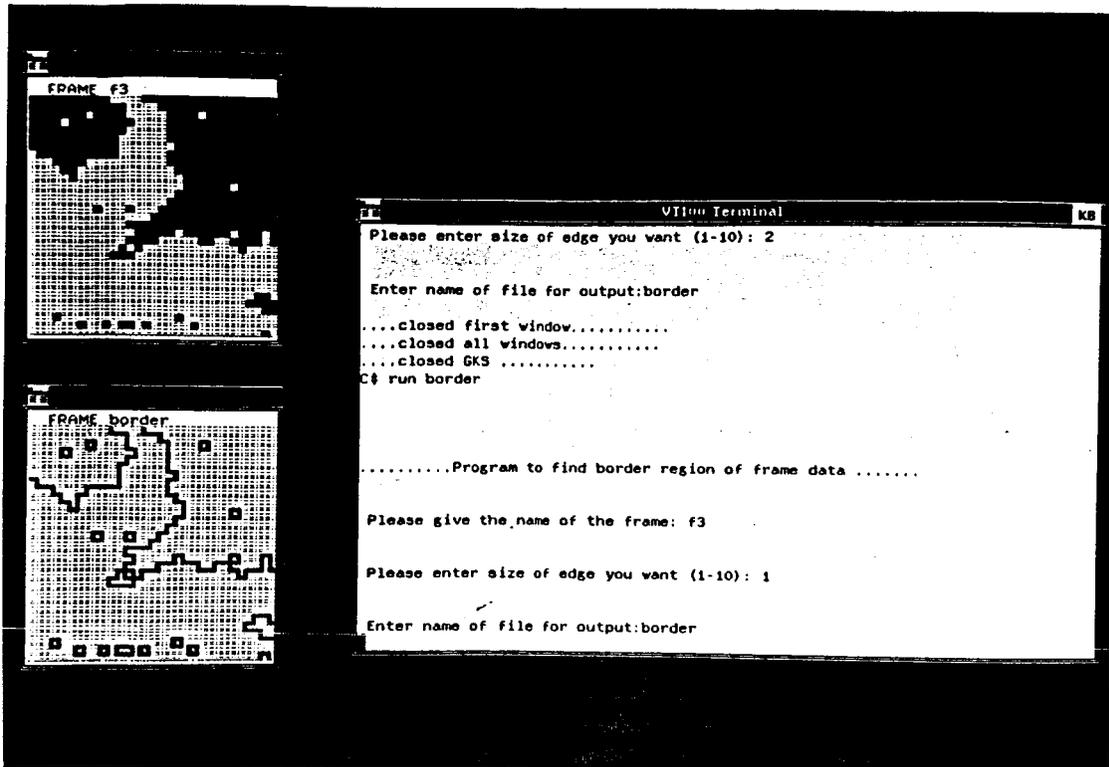


Figure 4

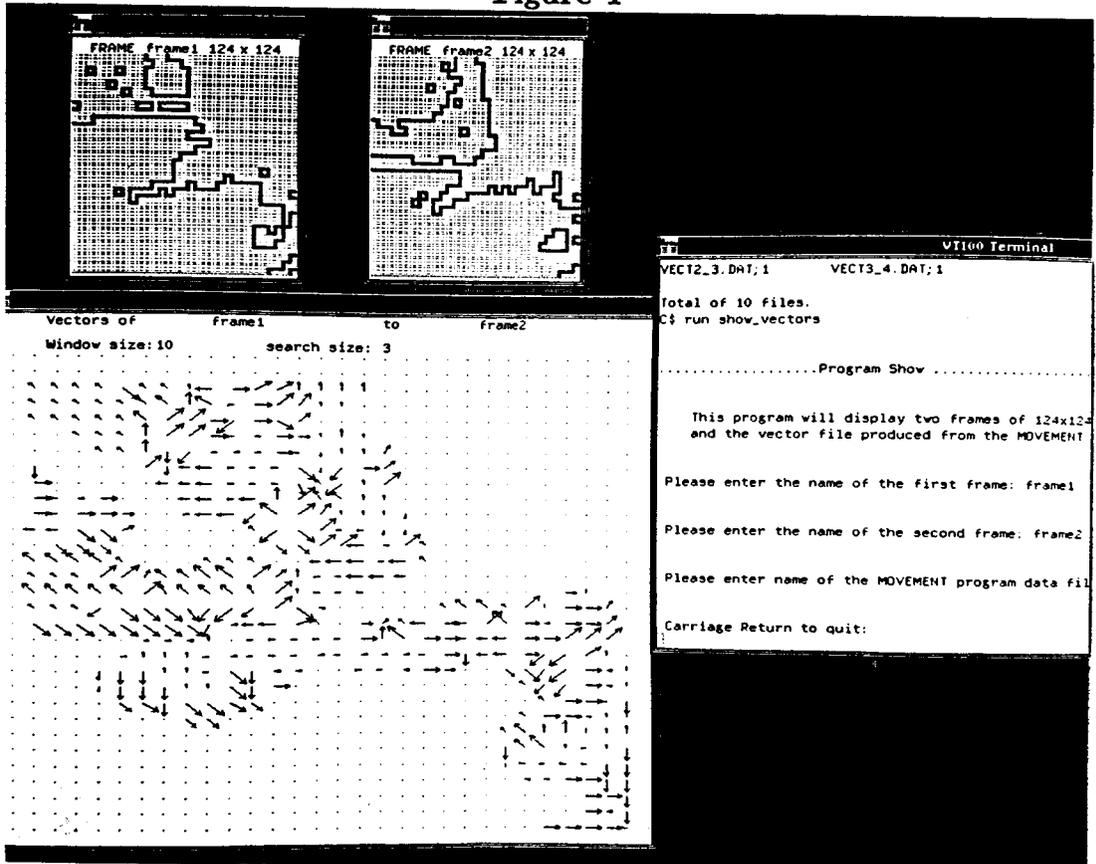


Figure 5a

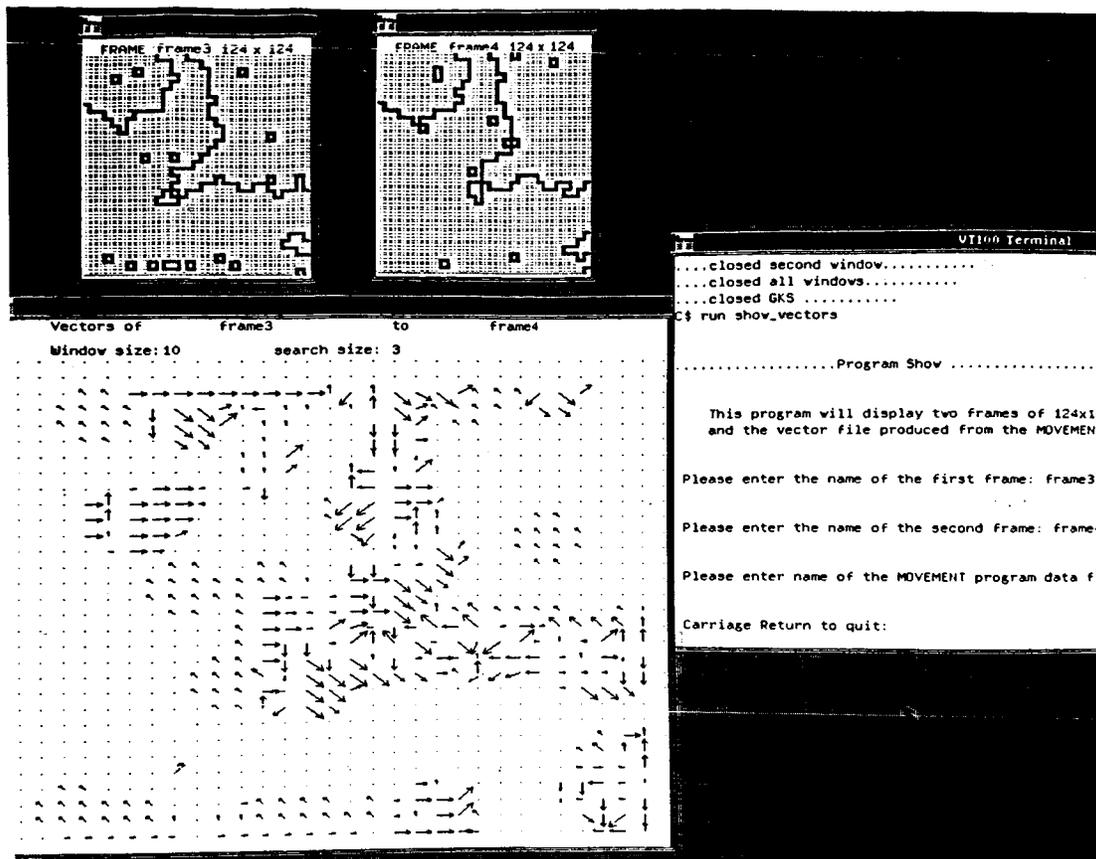


Figure 5b

ORIGINAL PAGE IS  
OF POOR QUALITY

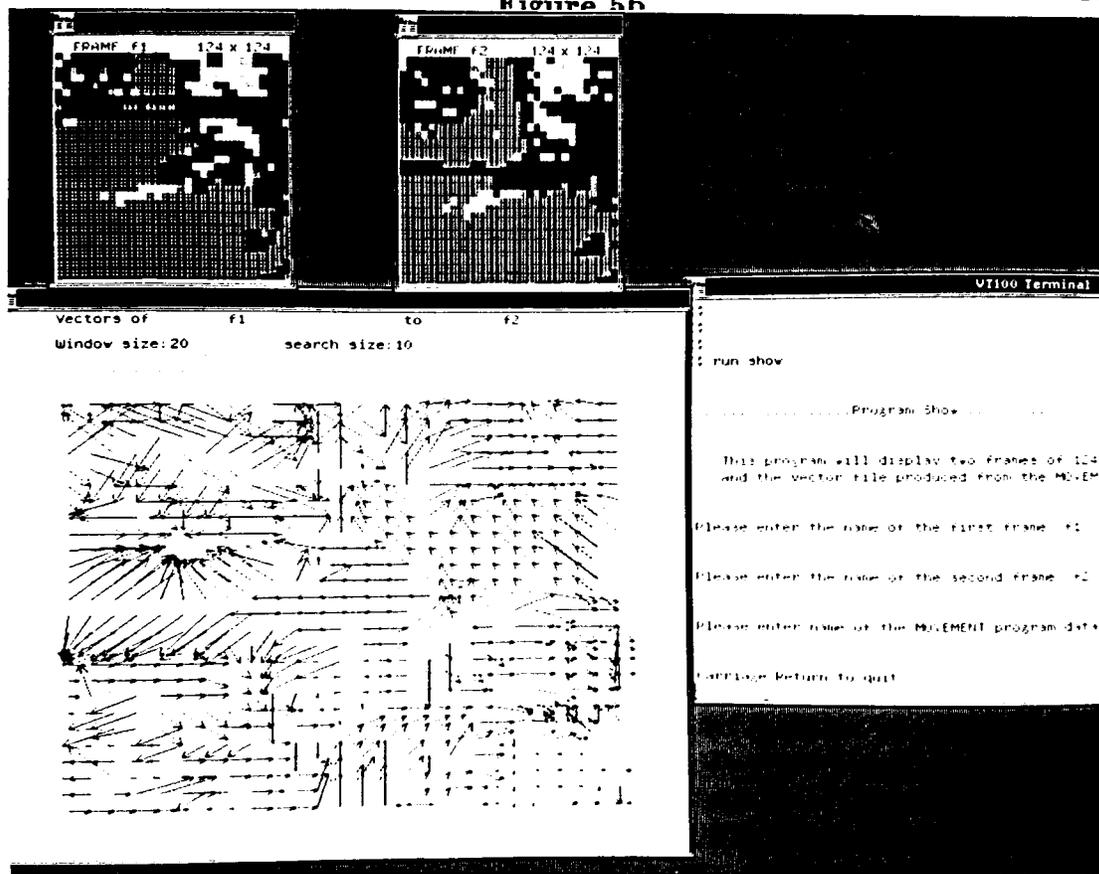


Figure 6

ORIGINAL PAGE IS  
OF POOR QUALITY

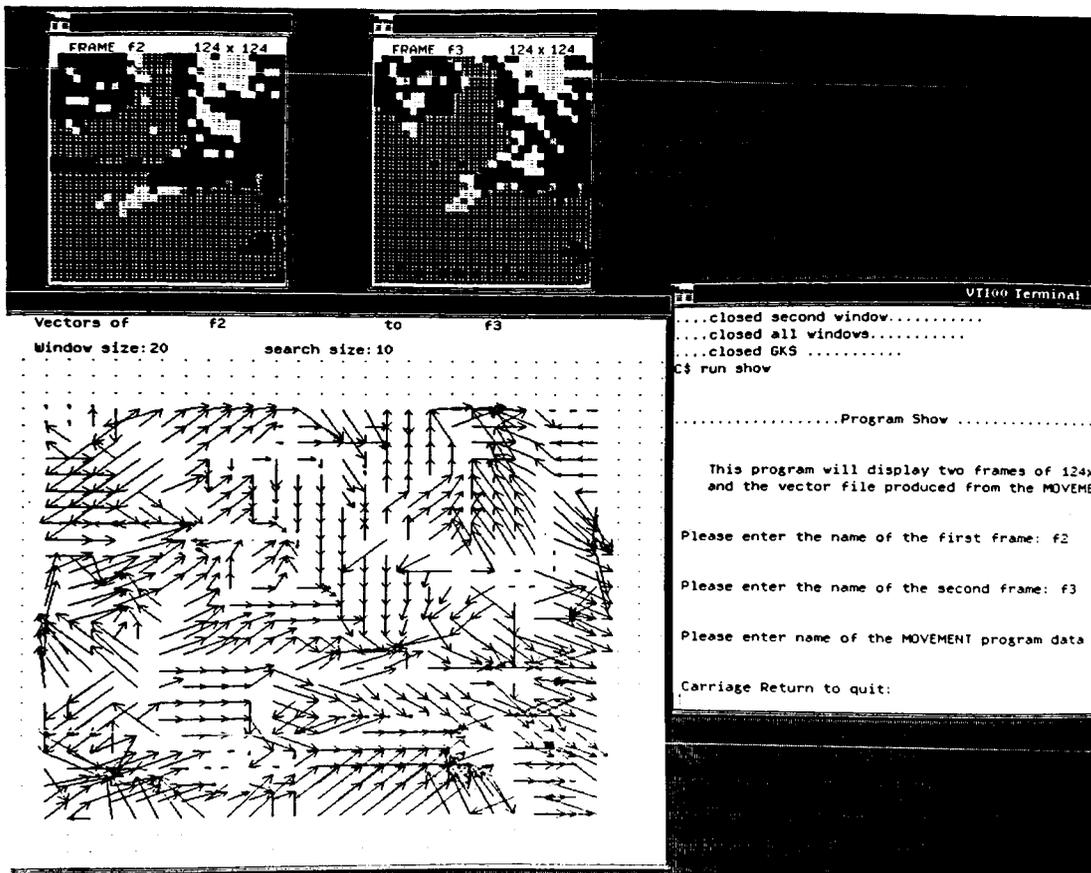


Figure 7

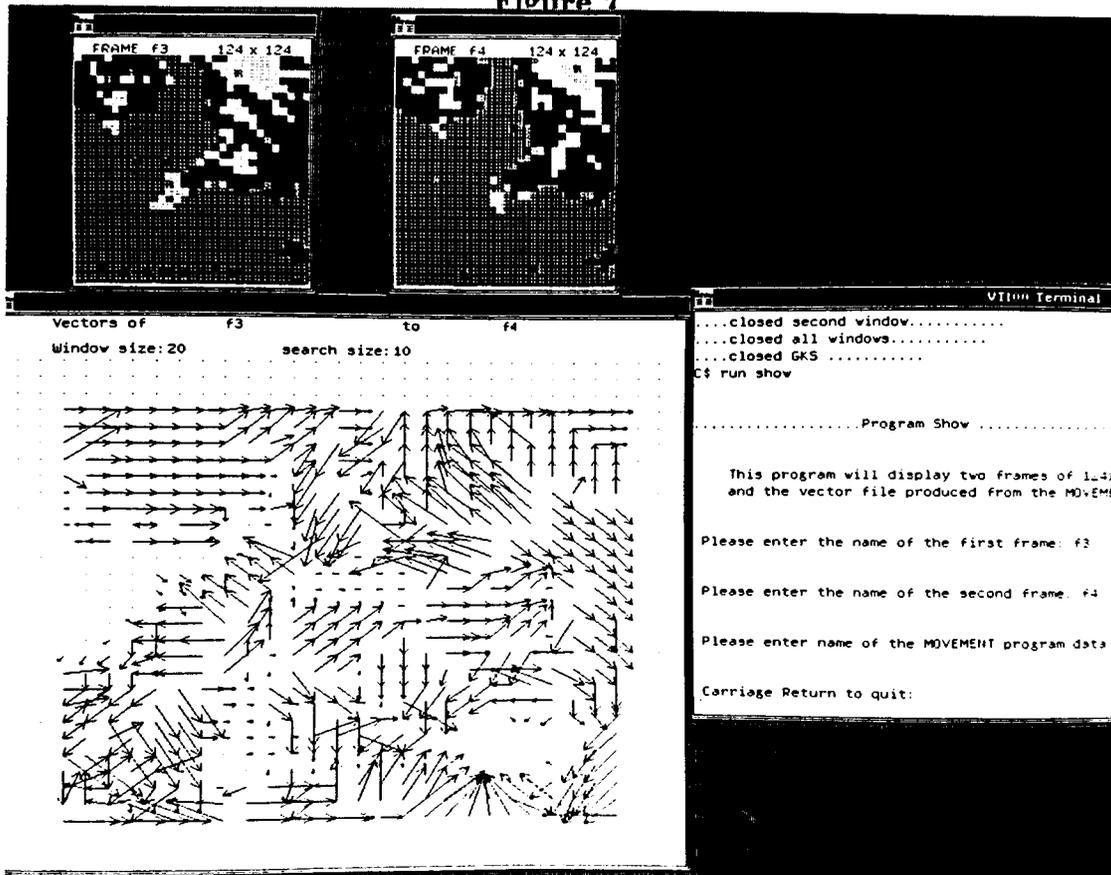


Figure 8

## APPENDIX (FROM THE PHASE I PROPOSAL)

### Identification and Significance of the Innovation

The overall objective of this proposal is to develop a method for extracting time derivative information from geostationary meteorological satellite imagery for the purpose of improving numerical weather prediction. In Phase I, we shall undertake to carry out a proof-of-concept study to demonstrate the feasibility of using pattern recognition techniques and a statistical cloud classification method to estimate time rates of change of large-scale meteorological fields from remote sensing data. Our intention in Phase II is to develop practical algorithms suitable for operational use and to implement these algorithms in an integrated stand-alone work station.

Great progress has been made in assimilating meteorological satellite data to aid in determining initial conditions for large-scale numerical weather prediction. Much of the recent increase in skill of operational forecasts is due to improvements in the initial state, brought about by modern objective analysis and initialization schemes which extract statistically optimal estimates of the initial condition from a heterogeneous data set. In this process, asynoptic satellite data are employed to augment the conventional surface-based network of synoptic observations. The result is two analyses per day, at noon and midnight Greenwich time, from which operational weather forecasts are produced numerically.

Geostationary Visible Infrared Spin-Scan Radiometer (VISSR) and VISSR Atmospheric Sounder (VAS) satellite images, however, are available at half-hour intervals and could be used to verify the evolution of the meteorological fields during the first time steps of the numerical forecast integration. Much of the information content of these images is not used in constructing initial states. If analyzed by pattern recognition techniques to extract their full information content, sequences of such images provide the potential for determining observational estimates of the time derivatives of meteorological fields during the first few hours of the large-scale forecast period. Forecast verification studies, in conjunction with predictability theories, show that these early hours of the prediction are typically characterized by a rapid decay of skill. If observational estimates of time derivatives can be compared with numerical computations during this critical period, a potentially valuable new tool will have been made available for improving numerical weather prediction models. Such improvements may take several forms. For example, it may prove practical to correct dynamical models statistically using this new type of observational information, as well as to investigate the physical causes for dynamical model errors and thus to improve the realism of the models themselves.

## **Phase I Technical Objectives**

The time rate of change of cloud features on visible or infrared geostationary satellite images separated by a short time (one half hour to one hour) can provide information on time derivatives of meteorologically significant fields, such as winds at cloud levels, intensification and movement of storms, and frontogenesis. In Phase I, we propose to investigate two cases: the short-term evolution of extratropical cyclone features and that of the mid-latitude marine atmospheric boundary layer. In the first case, the analysis technique should be able to recognize characteristic cyclone features automatically. In the second, it should permit the classification of cloud type as well. The cloud classification technique recently developed by Tournadre and Gautier (1985) and described below can handle both kinds of pattern recognition.

The specific technical objectives of Phase I are:

1. Develop and test a cloud classification technique based on pattern recognition methods, suitable for the analysis of visible and infrared geostationary satellite VISSR imagery.
2. Develop and test a methodology for intercomparing successive images using the cloud classification technique, so as to obtain estimates of the time rate of change of meteorological fields.
3. Implement this technique in a testbed system incorporating an interactive graphics terminal to demonstrate the feasibility of extracting time derivative information suitable for comparison with numerical weather prediction products.

## **Phase I Work Plan**

### *Overview*

We foresee that the crucial element in demonstrating the viability of our concept will be the successful development of a pattern recognition methodology which will permit the analysis of infrared and visible geostationary satellite imagery in a reliable, semi-automated fashion. This methodology must be able to identify key aspects of cloud features in the satellite images unambiguously, so that successive images, which may be separated in time by one half hour to one hour typically, can be intercompared to yield useful information on the time rate of change of synoptically significant meteorological fields. Once the pattern recognition techniques have been developed and refined, the next major task will be to implement software for applying them to the analysis of actual real-data images, using an interactive graphics terminal. This terminal will be driven by a VAX-11/730 supermini

computer system during the proposed research, but this configuration may be regarded as a testbed for the integrated work station to be developed in Phase II, which we envision will be based on a microcomputer system. Within Phase I, however, we shall conduct extensive numerical experimentation on the more versatile computer system (VAX-11/730) to determine the dependence of the results on parameters such as image size and resolution, as well as on the variable parameters of our pattern recognition scheme. This scheme itself is centered on a newly developed and versatile cloud classification scheme, and so we next devote the major part of this section to a brief description of that scheme.

### *Statistical Cloud Classification Method*

Before clouds can be classified, they must be characterized by a convenient number of parameters, the choice of which depends on the particular application. Following Parikh (1977), Weinman et al. (1985), and others, we use both spectral (visible, infrared and perhaps eventually microwave) and textural (mean of distribution, contrast, second angular momentum and entropy) parameters.

Once the set of parameters describing cloud field characteristics has been selected, satellite images can be analyzed to produce sets of these parameters for  $n \times n$  pixel boxes from an initial image, in a way similar to the production of sets of profiles from radiosonde measurements. The set of box parameters can then be classified using the Typical Shape Function (TSF) analysis technique (Jalickee and Ropolewski, 1979; McMillin, 1985).

### *Typical Shape Function Analysis*

The TSF analysis is used to classify, in a restrictive and distinguishable form, the set of parameters describing the cloud fields. The first step in this classification is to remove the mean value of each parameter in order to analyze only fluctuations. Then the matrix  $A$  of parameter sets for an entire image is decomposed into singular values:

$$A = X\Lambda Y^T$$

$X$  is the left singular vector matrix,  $Y$  is a right singular vector matrix and  $\Lambda$  is the diagonal matrix of the singular values. Only the eigenvalues required to explain 95% of the variance are retained. The associated eigenvectors, however, are in themselves insufficient to determine uniquely an exclusive classification. Such an exclusive classification can be obtained by computing a rotation matrix  $\alpha$  such that:

$$\hat{Y}^T = \alpha Y^T \quad \text{and}$$

$$\hat{A} = X\Lambda\beta^T, \quad \text{with } \beta^T = \alpha^{-1},$$

$\alpha$  is calculated iteratively by minimizing the rows of the right singular vector matrix  $Y$ , using Lagrange multipliers. The columns of  $\hat{A}$  then become the Typical Shape Functions (TSF). Once obtained, these TSF can be employed to divide the data set into classes by computing the scalar product of the parameter sets and the TSF.

### *Feature Recognition*

Cloud features and morphology can readily be recognized and outlined using this classification scheme. In the case of extratropical cyclones, the characteristics to be extracted and quantified include the frontal cloud mass extension, the precise cold front location, and possibly the cloud structure in the cold air mass. This is expected to be achievable with the technique outlined above, and the feasibility of recognizing these features will be investigated thoroughly during the Phase I period. The next step will be to develop methods for reliably tracking features from one image to the succeeding image, in a manner similar to that of Bouthemy and Benveniste (1984) or to methods used for deriving winds from cloud motions. Preliminary tests show that it is indeed possible to retrieve similar classes on two consecutive images separated by about one hour. Further tests are needed to establish the precision which can be achieved in the determination of the location and extension of the cloud mass, for example. In the case of the marine boundary layer, the advection of different air masses is characterized by different cloud types and morphologies. Because our proposed analysis technique is capable of distinguishing among several kinds of cloud cover, we expect to be able to compare cloud model predictions with cloud observations partitioned into classes.

### *Intercomparison with Numerical Weather Prediction*

Once we are able to estimate time rates of change (TROC) of meteorological variables from satellite imagery, we intend to carry out an illustrative comparison between a sample of our observation-based TROCs and a corresponding set of TROCs from an idealized numerical weather prediction model. For this purpose, we plan to use a quasi-geostrophic low-order spectral model (Roads and Somerville, 1984) initialized with conventional meteorological data. Such a model is economical to integrate but produces creditably skillful short-range predictions of the large-scale features of the evolving atmospheric circulation. Such a model contains a limited number of adjustable parameters representing the effects of subgrid physical processes, such as friction. To illustrate the potential utility of observational TROCs in improving dynamical forecast models, we plan to "tune" one or more such model parameters in regions of extratropical cyclone activity where model and observed fields of TROC differ systematically. Such a simple tuning in an idealized model is

a surrogate for the development of realistic parametric representations. If, as we expect, it proves possible to improve forecast skill significantly by such a simple approach, we anticipate testing more ambitious applications of our concept, using physically comprehensive state-of-the-art forecast models, in Phase II.

## Bibliography

The following are cited in the text of the proposal:

- Bouthemy, P. and A. Benveniste, 1983: Atmospheric Disturbances Tracking in Satellite Images, *Image Sequence Processing and Dynamic Scene Analysis*, T.S. Huang, eds., Springer-Verlag, Berlin.
- Jalickee, J.B. and C.F. Ropolewski, 1979: An Objective Analysis of the Boundary-Layer Thermodynamic Structure during GATE, *Monthly Weather Review*, 107, 68-76.
- McMillin, L.M., 1985: Use of Typical Shape Functions of Radiance Vectors in Satellite Retrievals, to appear in *Advances in Remote Sensing Retrieval Methods*, published by IFAORS, A. Deepak, Editor.
- Parikh, J., 1977: A Comparative Study of Cloud Classification Techniques, *Remote Sensing of Environment*, 6, 67-81.
- Roads, J.O., and R.C.J. Somerville, 1984: Linear predictability: Effects of stationary forcing. *Predictability of Fluid Motions*, G. Holloway and B.J. West (eds.), American Institute of Physics, 106 New York, 557-570.
- Tournadre J. and C. Gautier, 1985: Satellite Cloud Morphology Analysis using Typical Shape Functions, to be submitted to *Journal of Climate and Applied Meteorology*.
- Weinman, J.A., B.R. Markstrom and R. Wu, 1985: Applications of Computer Pattern Analysis Techniques to Meteorological Satellite Images-Rainfall Estimation and Pattern Analysis of Extra-Tropical Cyclones, to appear in *Advances in Remote Sensing Retrieval Methods*, published by IFAORS, A. Deepak, Editor.

## Additional References:

The following general references describe techniques on which we shall draw:

- Image Sequence Processing and Dynamic Scene Analysis, edited by T.S.Huang, NATO ASI Series, Springer-Verlag, 1983.

Parameter Estimation, H.W.Sorenson, Marcel Dekker, 1980.

Discrete Techniques of Parameter Estimation, J.M.Mendel, Marcel Dekker, 1973.

Statistical Estimation, Asymptotic Theory, I.A.Ibragimov, R.Z.Has'minskii, Springer-Verlag, 1981.

Estimation of Dependences Based on Empirical Data, VI.Vapnik (specific chapters on "Methods and Algorithms of Parametric Statistics for the Pattern Recognition Problem"), Springer-Verlag, 1982

Applied Time Series Analysis, R.K.Otnes, L.Enochson, John Wiley and Sons, 1978.

Image Pattern Recognition, V.A.Kovalevsky, Springer-Verlag, 1980.

### **Relationship with Future Research and R&D**

Phase I of this proposal, as described above, is a proof-of-concept study to demonstrate the feasibility of using pattern recognition techniques and a statistical cloud classification method to estimate time rates of change of large-scale meteorological fields from remote sensing data. The anticipated results are the successful achievement of the technical objectives listed above: development of the cloud classification technique, development of the intercomparison methodology for estimating time rates of change, and implementation in a testbed system incorporating an interactive graphics terminal. We also anticipate the successful testing of the concept of using these data to improve a numerical weather prediction model, as outlined above.

Phase II will build on the foundation of these results to produce a fully operational stand-alone work station to enable the benefits of this research to be realized in a real-time numerical weather prediction environment. We anticipate that this work station will include an interactive graphics terminal driven by a microprocessor. Software included with this work station will embody the features outlined above and will be implemented using artificial intelligence concepts to facilitate use in an operational forecasting center.