Report DARPA/SEMI/1089

OPTICAL NEURAL NETS FOR SCENE ANALYSIS

David Casasent (PI)
Carnegie Mellon University
Center for Excellence in Optical Data Processing
Department of Electrical and Computer Engineering
Pittsburgh, PA 15213
(412) 268-2464

DTIC
ELECTE
OCT.26 1989
B

23 October 1989

UNCLASSIFIED - DISTRIBUTION UNLIMITED

Date 23 Oct 89   D. Casasent(PI)

AD-A213 986

89 10 26 121

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	Unclassified

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
DARPA/Semi/1089	

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Carnegie Mellon University		U.S. Army Missile Command

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
Department of Electrical and Computer Engineering Pittsburgh, PA 15213	AMSMI-RD-DP-ST Redstone Arsenal, AL 35898-5244

8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Defense Advanced Research Projects Agency		Contract DAAH01-89-C-0418

8c. ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
1400 Wilson Blvd. Arlington, VA 22209				

11. TITLE (Include Security Classification)

"Optical Neural Nets for Scene Analysis" (Unclassified)

12. PERSONAL AUTHOR(S)
David Casasent

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Semiannual Technical	FROM 04/1/89 TO 10/1/89	18 October 1989	10

16. SUPPLEMENTARY NOTATION

17	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Optical neural nets, Optimization neural nets, Production system neural net, Adaptive clustering neural net, Matrix inversion neural net, Symbolic neural net

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

A hybrid optical/digital neural net for scene analysis is described. It combines pattern recognition and neural net techniques. New algorithms, architectures and applications are described for optimization neural nets (a mixture neural net for image spectrometry, cubic and quadratic neural nets for multitarget tracking, and a matrix inversion neural net), production system neural nets, symbolic neural nets and a new adaptive learning neural net (the adaptive clustering neural net). Progress in the first six months on these seven neural nets and our hardware are presented. Keywords for Artificial Inte...

Computer Programming ...

Data ...

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Frederick Clarke	(205) 876-5987	AMSMI-RD-DP-ST

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

Semiannual Report (April-October 1989)
18 October 1989

OPTICAL NEURAL NETS FOR SCENE ANALYSIS

1. PROGRAM OVERVIEW/OBJECTIVE

The objective of this program is to develop new neural net (NN) algorithms and architectures for scene analysis.

2. APPROACHES

We have three general approaches. First, we distinguish between optimization and adaptive learning neural nets. Second, we attempt to combine advanced pattern recognition concepts and NN techniques into hybrid pattern recognition/NN architectures. Third, we address new hybrid optical/digital NN algorithms and architectures. We now briefly discuss these approaches.

There are various levels of scene analysis that must be appreciated [1]. Scene analysis has many unique problems that do not arise in other data processing problems. These include: shift-invariance, distortion-invariance feature extraction, segmentation, and classification. The input neuron representation space used is also of significant concern, since the use of iconic (pixel) inputs requires many neurons (250,000 for a 500x500 pixel input). We achieve shift and limited distortion-invariance by the use of a feature space input. This also appreciably reduces the number of input neurons required. It also reduces the amount of training required (since the system need not be trained on all distorted versions of each object). We assume that the feature space is optically produced by another processor. To achieve segmentation (isolation of one object in the field of view) we use range imagery and a symbolic correlator. The NN achieves classification of the input object. During learning it modifies the initial feature space and provides more complete distortion invariance.

We will use a new adaptive clustering NN (ACNN) for our adaptive learning NN. The ACNN combines pattern recognition techniques (linear discriminant functions, clustering and feature spaces) and NN techniques (to achieve piecewise linear discriminant surfaces). We consider three optimization NNs: (1) a cubic energy and (2) a quadratic energy MTT NN (these solve multiple constraint problems with the specific application considered being (MTT) multitarget tracking), and (3) a mixture NN (this solves a constrained least squares problem for the fractional amounts of different elements present in each image region with the application considered being the processing of imaging spectrometer multispectral data). The final three NNs we consider are: (4) a matrix inversion NN (this is most useful for all advanced linear algebra functions in image processing), (5) a production system NN (for performing if-then propositional and predicate calculus decisions and inferences), and (6) a symbolic NN (this combines a symbolic correlator and a production system and allows multiple objects to be handled). The ACNN (7) is the major NN we will consider.

We will be fabricating an optical laboratory neural net for on-line processing of input and hidden layer neuron levels and their weights. This will use new modulated error diffusion optical computer generated hologram (CGH) interconnection techniques. This optical NN will be used in

conjunction with a hardware Hecht-Nielson Corporation (HNC) digital neural net. Our initial hybrid optical/digital NN concept will involve the use of the digital NN for adaptive learning and the optical NN for classification with real-time processing (with its interconnection weights downloaded from the digital NN). Future goals involve increasing the use of adaptive functions on the optical NN as optical materials and components mature.

3. SUMMARY OF PROPOSED NNs

The seven neural nets we will consider are now briefly summarized.

The input neurons to the <u>production system NN</u> are facts (antecedents and consequents). Objects and object parts will be initially used and then surface types for object parts (cylinder, sphere, valley, ridge, etc.) will be used. The objects will be typical of those present in various scenes. The weights define the rules. These will initially be posed as if-then statements, with all rules written as the AND of several antecedents and the OR of several such sets of antecedents. The output neurons that fire represent the new facts that are now learned to be true. As the system iterates, it learns new rules and infers new results on the present input data. We initially consider a propositional calculus system (with all parameters being exact terms) and then will address a predicate calculus system (with parameters being variables) that is much more powerful.

The E input neurons in our <u>mixture NN</u> each correspond to the fractional amounts of E elements present in a mixture of elements within one region of an input scene. The outputs from two matrix-vector multiplications are combined to form the new neuron states. After a number of iterations, the final neuron states denote the fractional amount of each element present in the input mixture.

The <u>matrix inversion NN</u> produces the inverse of a matrix that is given to the processor. To calculate the inverse $\underline{X}$ of a matrix $\underline{Q}$, we realize that $\underline{QX} = \underline{I}$. We formulate the solution (the elements of the inverse of Q with elements) as the minimization of an energy function . We solve for the X that minimizes the energy function on a neural net. The matrix elements (weights) in this NN have an attractive block Toeplitz form and thus acousto-optic (AO) architectures should be very suitable for implementing this NN. This would represent the first AO NN. Since matrix inversions are required in many pattern recognition linear discriminant function designs and in most adaptive algorithms, this NN should have general computational use in image processing (as well as in adaptive radar, control, etc.).

The <u>cubic energy NN</u> for MTT takes measurements on objects in each of three frames and it assigns one target per measurement and time frame. This is useful for time sequential scene analysis to associate objects (or object parts) in several time frames.

The <u>quadratic energy MTT NN</u> is a simplified version of the cubic energy NN. It processes pairs of time frames. The resultant optical architecture will be much simpler and has significantly reduced component requirements.

The <u>symbolic NN</u> combines a symbolic correlator, production system NN, feature extractor and image processing NN. Its major advantage is the ability to process multiple objects in the

field of view (this is achieved by the symbolic correlator). It outputs a symbolic description of each region of the input that denotes which generic shapes are present and their location. These data are then symbolically encoded and fed to an NN. The NN is unique because of its symbolic input neuron representation. Alternatively, the locations of regions of interest in the input scene are used to guide the positioning of window functions (for segmentation) from which input features are extracted and subsequently fed to an NN for object classification. These NNs again combine pattern recognition and NN techniques.

The <u>adaptive clustering NN</u> will receive major attention. The input neurons will be features, the hidden layer neurons will be prototypes of the various classes of objects and the output neurons will denote the class of the input object. Clustering techniques will be used to select the original hidden layer neurons (we allow several neurons or clusters per object class) and hence the initial input to hidden layer weights. These represent a set of linear discriminant functions (LDFs). The output neurons define the class of the input. The hidden to output layer weights map the clusters to classes. Our study of criterion functions determines the type of error function used to train the NN. Thus, advanced pattern recognition techniques are used to initialize the set of NN weights. A new adaptive NN learning algorithm is then used to refine and improve the initial weight estimates and to provide the LDF combinations that provide the nonlinear piecewise discriminant surfaces finally used. This is the adaptive learning stage. This new NN combines pattern recognition and NN techniques.

4. YEAR 1 PLANS

For year 1 (April 1989-March 1990), we have isolated 4 tasks that the man years available allow.

- **Task 1:** Fabrication of the initial optical lab NN and interfacing of the Hecht-Nielson Corporation (HNC) digital NN.

- **Task 2:** Completion of our MTT (cubic and quadratic) NN studies and simulations.

- **Task 3:** Simulations of the initial production system NN on object part input data.

- **Task 4:** Formulation of the first level of our adaptive clustering NN and simulated demonstrations of its use and performance.

The personnel involved include: David Casasent (principal investigator), Frank Matousek (technician), Sanjay Natarajan (neural net hardware, ACNN, imaging spectrometer), John-Scott Smokelin (matrix inversion neural net), and Mark Yee (MTT neural net and production system).

5. SUMMARY OF PROGRESS (MONTHS 1-6)

Our work has been published in Journal and Conference papers and thus we reference these papers and now summarize our work in the first 6 months. This has been quite prolific as we have detailed the concepts for 7 different neural nets. We now summarize our work on these 7 NNs and our laboratory hardware.

5.1 Production System Neural Net

The basic concept for this new type of NN was detailed [2]. We considered several realizations of it and find a NN one preferable. We formulate all rules as IF-THEN statements with antecedents a_n on the left and consequents c_n on the right. We allow the AND of several antecedents and the OR of several such combinations. As our new input neuron representation space, we use one neuron per fact (antecedent or consequent). The input neurons are binary and are activated "1" if a fact is true or inactive "0" otherwise. After several iterations, the system learns all consequents. The system can learn new rules, not explicitly present in the original IF-THEN rules. The system can also operate with analog neurons, whose activation level is proportional to the confidence of antecedent facts (as probabilities). We use analog weights and binarize the output neurons. We formulated the weights as a matrix and the input and output neurons (facts) as vectors. This allows implementation of this NN on an optical matrix vector processor. The rules and hence the matrices in the optical matrix-vector realization of this NN are fixed. Hence, its optical realization is most attractive. This represents a very new use of an optical matrix-vector processor and a NN, with facts being the input neuron representation space used.

5.2 Mixture Neural Net

The application we consider for this NN is the processing of imaging spectrometer data. This is a specific case of a general mixture problem in which the amount x_e of each of E elements present in an input mixture signal $\underline{c}$ must be determined. Generally only several elements are present out of perhaps 500. This analysis must be performed for every pixel in an input scene to identify what is present, interesting, new, etc. We have described this problem as an NN and detailed its optical realization [3]. The optical realization uses two matrix-vector multipliers, whose outputs are summed, thresholded, and fed back to the input of one of the matrix-vector multipliers. The final vector output denotes the x_e. The matrices in this optimization NN are fixed (the database of element spectra, etc.) and thus its optical realization is very attractive. This mixture NN represents a new NN applications.

5.3 Matrix-Inversion Neural Net

This matrix-inversion NN represents one of the most general purpose NN applications. This occurs since matrix-inversion is a basic operation in all of modern signal and data processing, which emphasizes operation on data arrays. Specific examples occur in pattern recognition, adaptive signal processing, adaptive phased array radar, etc. The basic idea was outline in Section 3. It has been more fully detailed [4], several new acousto-optic (AP) neural net architectures to implement it have been advanced, and initial simulation results have been obtained and are most promising. This is the first AO NN and it can easily handle NN problems larger than the dimensionality (space bandwidth product) of the AO cell; plus, it is one of the most general NNs we consider.

5.4 Cubic Energy Neural Net for Multitarget Tracking (MTT)

We have detailed the cubic energy NN for a specific multitarget tracking (MTT) problem [5]. This has general use in time-sequential image processing. For MTT, the goal is to associate measurements in a sequence of frames with different targets. We have detailed the algorithm for the case of 3 time-sequential frames of data. We have devised a new optical architecture to

achieve this. New encoding, multiplexing, architectural and algorithmic concepts were used to significantly reduce the component requirements. Since this is a tensor rather than a matrix problem, such issues arise. Despite these improvements, we found the present state of optical components too primitive to pursue this architecture further. Thus, we developed the quadratic MTT NN (Section 5.5). Our contributions in terms of architectures and algorithms for the cubic NN are still very significant and are suitable for all tensor (versus matrix-vector) NN problems. Thus, this completed effort should have useful future impact as various nonlinear and photorefractive optical components mature.

5.5 Quadratic Energy MTT Neural Net

This new and simplified algorithm [6] is far more attractive and is more easily realized with more easily foreseen optical technology. Thus, we can conceive of it being realized in optical hardware with a concentrated 4 year effort starting in 1990. Presently, we have demonstrated its basic concept and its ability to handle multiple objects in a field of view [6]. We have also detailed its use in satellite rendezvous and docking [7] with promising initial results (under a NASA grant). These results are most attractive when noise is present. Extensions are presently being pursued for this quadratic energy MTT NN.

5.6 Symbolic NN

This symbolic NN concept is unique as it offers the only presently viable solution to the processing of multiple objects in the field of view (FOV) of a 2-D image sensor. Our present version consists of a symbolic correlator [8] whose outputs feed a production system NN [2]. The symbolic correlator allows the use of this system with multiple objects in the FOV. It also allows for high capacity and distortion invariance. Since it is a correlator, it achieves shift invariance and can handle multiple objects. These symbolic outputs from each region of the FOV can encode a large number of objects [9]. The unification of this symbolic correlator and production system NN into a symbolic NN will be provided soon [10] with examples and simulated data for a specific case study. This new NN offers the only viable NN solution to multiple objects in the FOV of a sensor. This concept of a multichannel optical correlator (using smart filters and symbolic encoding) feeding a production system NN (to achieve a symbolic NN) is most unique and powerful.

5.7 Adaptive Clustering NN (ACNN)

This represents one of our projects that will run all 3 years. It was highlighted in Section 3 and epitomizes our approach (Section 2) to a hybrid optical/digital adaptive learning NN and a hybrid pattern recognition/NN classifier for scene analysis. It uses a 3 layer NN, which can achieve any piecewise nonlinear discriminant surface. We consider only supervised learning NNs.

We started the study of this ACNN with a review of criteria (error) functions used in standard pattern recognition. We showed that the perceptron and sigmoid criterion functions are both useful over a wide range of parameters [11]. We find the perceptron criterion function to be most useful (since the minimum energy solution can be reached more easily with it than with the more standard sigmoid criterion function). This represents a significant departure from standard NNs. We note that the thresholding (nonlinear) function used in NN classification (we distinguish between training and classification) can and should be different than the most popular sigmoid function [11].

This review of criterion functions [11] and our approaches (Section 2) lead us to our new ACNN concept. It was first advanced [1,11] and recently fully detailed [12]. We note that another NN contribution in these papers [11,12] is that conjugate gradient techniques are preferable (in all NNs) to the more conventional delta or gradient descent methods.

We also observed that the hidden layer neurons are often binary in the most popular adaptive learning NN backpropagation (BP) algorithm [11]. We generalized this to assign specific initial hidden layer neurons to clusters or prototypes that are representative of the classes in the input data. Thus, we employ pattern recognition clustering techniques to determine the number of hidden layer neurons to use and to produce the initial set of input to hidden layer neuron weights.

Use of such an initial set of weights is found to be preferable to the random starting weights used in BP etc. As learning proceeds, we adapt these weights using a new algorithm [12]. In classification, we perform maximum selection of the most active hidden layer neuron. Thus, our hidden layer neurons are binary. The hidden-to-output layer neuron connections are also binary and perform mapping from the prototypes (hidden layer neuron selected) to the output neurons (that denote the class of the input data).

Thus, in our 3 layer ACNN, the input neurons are analog and represent distortion-invariant features, the hidden layer neurons are binary and are prototypes or clusters, while the binary output neurons denote the class of the input data. As our input neuron representation space, we use a feature space (wedge-sampled Fourier coefficients) that allows some in-plane distortion invariance. This pattern recognition technique also significantly reduces the number of input neurons and interconnections required (since this neuron feature space is of reduced dimensionality compared to the iconic input image space). It also reduces the amount of training needed (since we need not show the NN all distorted versions of each object). Our ACNN also has the property that the number of parameters to be chosen empirically is significantly reduced. For example, our clustering technique determines the number of hidden layer neurons, our training algorithm does not require choices for momentum etc. parameters (as in BP), our conjugate gradient algorithm has no convergence and step size parameters to be chosen (as in gradient descent).

We have tested our ACNN on two databases: a synthetic two-feature 3-class set and a 3-D distorted set of aspect views of different aircraft (using a wedge-sampled Fourier feature space). The results obtained [12] are attractive. The ACNN trains much faster than BP and converges faster. Our new conjugate gradient technique significantly improves the training time for BP, and can be applied to other NNs. Our initial cluster or prototype selection adds little to the overall calculations required and reduces learning time (since we start from initial prototype weights, rather than from arbitrary weights as done in BP). The ACNN and BP achieve comparable performance with much less training time required for our ACNN. The combination of pattern recognition and NN techniques that our ACNN offers are most attractive. Its off-line digital training and on-line optical classification result in an attractive hybrid optical/digital NN. Our hardware for this system will be detailed in the next quarter.

6. PLANS FOR THE REMAINDER OF YEAR 1

We will document our new hybrid optical/digital neural net concept and initial hardware [13] and will discuss [14] how this one hybrid neural net can be used for all 7 neural net realizations and as an associative memory/processor (the associative memory use is not DARPA supported). This hardware task will continue all 3 years.

We will detail simulated [10] and initial optical laboratory [15] results for our production system and our symbolic neural net. We will use a set of objects composed of different generic parts to demonstrate this.

We will perform the first error source analysis of any neural net [13]. Our mixture neural net will be used as the case study. The general modeling and techniques used can be applied to other NN applications. This thus represents a major new result useful for all optical NNs (as well as for digital NNs).

We will be addressing our matrix-inversion NN further with attention to the many applications of it and to the accuracy required. Our NN error source modeling work will be of use here. The accuracy requirements will be of major concern. New algorithms to reduce these issues can be developed (year 2) and applied to specific applications (year 3). This year 2 and 3 schedule is tentative.

Our cubic energy NN work has been concluded. Our quadratic energy NN work will be continued for the same MTT problem, with attention to algorithm extensions, crossing targets, more targets in the FOV, improved algorithms, etc. and these results will be documented [16].

We will continue our symbolic NN work (this presently incorporates a symbolic correlator and production system NN). Simulation results [10] and initial optical laboratory results [15] will be obtained.

Our ACNN work will involve including its adaptive learning algorithm in software in our digital NN and attention to a new feature space input [17] and database.

REFERENCES

1. E. Barnard and D. Casasent, "Image Processing for Image Understanding with Neural Nets", *Proc. IEEE International Joint Conference on Neural Networks (IJCNN)*, June 1989, Washington, D.C., Vol. 1, pp. I-111-115.

2. E. Botha, D. Casasent and E. Barnard, "Optical Production Systems Using Neural Networks and Symbolic Substitution", Applied Optics, Vol. 27, pp. 5185-5193, 15 December 1988.

3. E. Barnard and D. Casasent, "An Optical Neural Net for Classifying Image-Spectrometer Data", Applied Optics, Vol. 28, pp. 3129-3133, 1 August 1989.

4. E. Barnard and D. Casasent, "Optical Neural Net for Matrix Inversion", Applied Optics, Vol. 28, pp. 2499-2504, 1 July 1989.

5. E. Barnard and D. Casasent, "Multitarget Tracking with Cubic Energy Optical Neural Nets", Applied Optics, Vol. 28, pp. 791-798, 15 February 1989.

6. M. Yee, E. Barnard and D. Casasent, "Multitarget Tracking with an Optical Neural Net Using a Quadratic Energy Function", Proc. SPIE, Vol. 1192, November 1989.

7. A.P. Lee, D. Casasent, P. Vermeulen and E. Barnard, "Estimating Satellite and Motion Parameters using a Novelty Filter and Neural Net Tracker", Proc. SPIE, Vol. 1053, January 1989.

8. D. Casasent, "Optical AI Symbolic Correlators: Architecture and Filter Considerations", Proc. SPIE, Vol. 625, pp. 220-225, January 1986.

9. D. Casasent, "Unified Synthetic Discriminant Function Computational Formulation", Applied Optics, 23, pp. 1620-1627, May 1984.

10. D. Casasent and E. Botha, "A Symbolic Neural Net Production System: Obstacle Avoidance, Navigation, Shift-Invariance and Multiple Objects", Proc. SPIE, Vol. 1192, November 1989.

11. E. Barnard and D. Casasent, "A Comparison between Criterion Functions for Linear Classifiers, with an Application to Neural Nets", IEEE Trans. Systems, Man and Cybernetics, Vol. 19, No. 6, accepted for publication, November 1989.

12. D. Casasent and E. Barnard, "An Adaptive-Clustering Optical Neural Net", Applied Optics, submitted August 1989.

13. D. Casasent and T. Slagle, "A Hybrid Optical/Digital Neural Network", Proc. SPIE, January 1990, to be submitted.

14. D. Casasent, "A Multifunctional Hybrid Optical/Digital Neural Net", Proc. SPIE, April 1990, to be submitted.

15. David Casasent, Elizabeth Botha, Jin Yun Wang and Ren Chao Ye, "Optical laboratory realization of a symbolic production system", Proc. SPIE, April 1990, to be submitted.

16. D. Casasent and M. Yee, "Measurement-Based Neural Net Multi-Target Tracker", Proc. SPIE, April 1990, to be submitted.

17. E. Barnard and D. Casasent, "Trajectories and Constraint-Based Distortion-Invariance", Proc. SPIE, April 1990, to be submitted.