An Efficient Neural Network Architecture for Recognition of Spatial Pattern Invariants

Michael J. Healy and Thomas P. Caudell
Research and Technology
Computer Services Division
The Boeing Company
PO Box 24346
Mail Stop 7L-22
Seattle, Washington 98124-0346

We describe a neural architecture for efficient recognition of invariant features placed arbitrarily in patterns of data. The architecture provides versatility in invariant selection with minimal computation and storage requirements. Operating in "dumb" mode, the architecture, called the Big, Dumb Mass Detector (BDMD), autonomously extracts fixed-size subsets of input pattern components based upon total activation ("mass") in the receptive fields of detector nodes. The BDMD sends these subset patterns to a pattern recognition neural network (the PR network) whose input field has exactly as many nodes as the fixed-size subsets. Operating in "smart" mode, the BDMD can be driven by another part of the recognition system to selectively scan an input pattern. Among the capabilities of the BDMD are its ability to "zoom in" on a pattern feature of particular interest. Only the relatively few connections and associated computations required for recognizing single features need be implemented in the PR network.

1 Introduction

Recognizing spatial pattern features invariant to translation, rotation, and scale is a major problem in pattern recognition with vision, infrared, and other imaging sensors. We have developed a neural architecture that extracts invariant features arbitrarily located in patterns of data. It provides invariance to multiple feature displacements and orientations using a minimal amount of computation and storage.

The architecture obtains its efficiency through a separation of the two tasks of feature extraction and feature (or object) recognition (see Fig. 1). The extraction component is a neural network called the Big, Dumb Mass Detector (BDMD). The BDMD autonomously extracts a subset of input pattern components by sampling a fixed number of nodes in the input field through any of a set of pre-defined subset selection masks. The current subset selection is based upon competition in a hierarchy of detector nodes, each representing the activity over a mass of input nodes. Each input node's activation level represents an input pattern component. The BDMD sends each sampled subset to a pattern recognition (PR) neural network whose input field has exactly as many nodes as the fixed size of the masks. If the input patterns represent two-dimensional images, then image features with arbitrary locations, seen at arbitrary scales and rotations, appear centered at a standard scale and rotation in the input field of the recognition network. Only the nodes and connections required for recognizing single features, which are assumed to occupy few input nodes relative to the input pattern, need be implemented in the PR network.

2 The BDMD Architecture

The overall layout of a BDMD/PR network architecture is shown in Fig. 1, with greater detail shown in Figures 2 and 3. A variety of neural networks or other systems can be employed in the PR system; in the systems we have designed, the PR system is based upon adaptive resonance theory

0-7803-0559-0/92 $3.00 © 1992 IEEE

(ART) (for example, see [1]), and the BDMD itself employs some of the principles of ART.

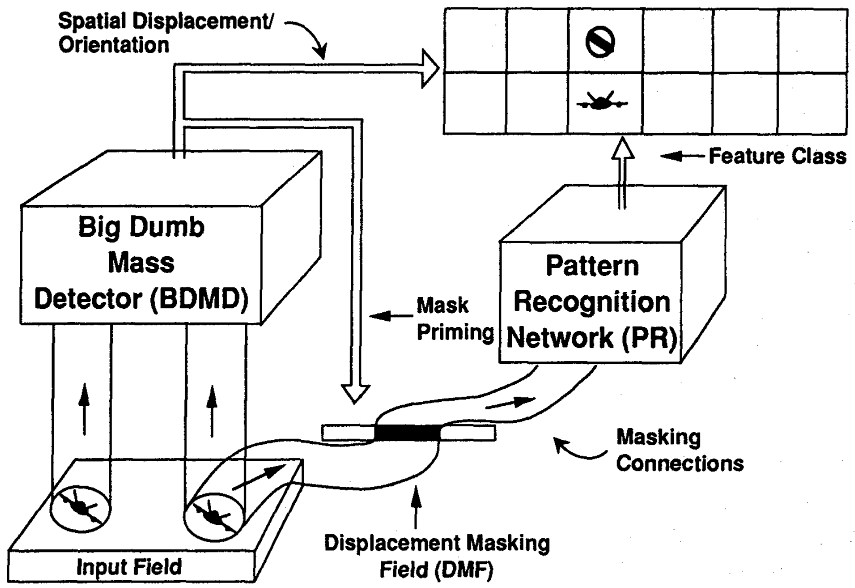

Figure 1. The BDMD/PR system employs a hierarchy of detector nodes, which respond to mass activity in their receptive fields and prime DMF masks.

The BDMD is based upon a multilayer hierarchy, with each layer corresponding to a single scale for feature extraction from the input field. Each node within the hierarchy is called a *detector node.* It samples a subset of input field nodes, called its *receptive field,* and competes with the other detector nodes through a system of inhibitory interconnects. The spatial scale of a receptive field corresponds to the layer containing the corresponding detector node, with scales increasing in the order bottom layer (D_1) to top layer (D_n).

The detector nodes control feature extraction by partially activating subsets of a *displacement masking field* (DMF). The partial activation is referred to as *priming.* As shown in Fig. 2, each DMF node receives input from a single input field node in addition to that from the detector nodes. Each DMF subset controlled by a detector node samples the receptive field of its detector node. All such subsets contain the same number of nodes, regardless of the size of the receptive field: This is accomplished through receptive field subsampling (Alternatively, it can be accomplished by averaging the activation values of several receptive field nodes per DMF node, but this requires multiple DMFs). A single DMF node samples several overlapping receptive fields at a variety of scales (Fig. 4), hence, is controlled by several detector nodes at different levels D_i. Each DMF subset controlled by a detector node possesses one-to-one output connections to a *displacement-free field* (DFF). The nodes of the DFF, in turn, possess one-to-one output connections to the input nodes of either the PR system, or an intervening rotation detection system as shown. Each DMF node, once primed by a detector node, produces an output signal value equal to the value of its corresponding node in the input field, which lies in the receptive field of the detector node. Thus, since the detector node receptive fields cover different regions of the input field at a variety of scales, the DMF transmits spatially-invariant pattern features to the DFF.

Rotation invariance for two-dimensional image features is handled in a similar manner, by a separate subsystem called a *rotation detector* (RD). In this scheme, the PR system samples the DFF through an intervening *rotation masking field* (RMF), whose purpose is similar to that of the DMF. Subsets of the RMF nodes are activated by the RD detector nodes. Each RD detector node samples the entire DFF, but through the intervening input layer of the RD system. The RD system is an ART system [1] with fixed templates: These have the form of bar-shaped patterns, each at a different rotation as indicated in Fig. 3. An RD detector node will be selected by the ART RD system to represent the current DFF pattern based upon resonance between its bar template and the DFF pattern.

A fundamental design problem in this architecture is that the PR can only recognize features if they are presented through the masking connections one at a time. Also, a collection of features, extracted and recognized sequentially, must be accumulated and organized in some fashion to provide useful information about an input pattern. This information includes the occurrence of multiple objects, appearing as multiple features of types familiar to the PR. These two considerations are addressed in Sections 3 and 4, respectively.

3 System Operation

An input pattern appears as either a binary (0, 1) or an analog (0, 0.33, 0.5, 0.872, ...) pattern of activation across the input nodes. For simplicity, we shall assume that the pattern remains constant for a sufficiently long time for it to be processed by the BDMD/PR system as we are about to describe. The processing involves the successive extraction and recognition of several features. Separate connections emanating from detector nodes, indicated in Fig. 2, retain the information about the original displacement of the extracted feature. The recognition output of the PR system, together with associated displacements and rotations from the BDMD and RD, forms one component of a multiple-feature pattern in *working memory.* Working memory is a system of nodes each of which can remain in an excited state long enough for the entire system to accumulate several input pattern features through the BDMD/PR system.

Feature extraction is facilitated by the multilayer system of detector nodes under the control of a reset node, as shown in Fig. 2. Detector nodes are so named because they compete to determine which receptive field will be read by the PR subsystem based upon *mass detection*, where mass is defined as the total amount of activation over a receptive field. Initially, the activation level of a detector node is an increasing function of the summed, excitatory connection-weighted input from its receptive field. Once excited, the detector nodes compete in a winner-take-all fashion through inhibitory interconnects (not shown), the winning node being one with maximal excitatory input. The winning detector node suppresses the other detector nodes while priming its subset of the DMF. DMF nodes possess the property that they must receive input from both the detector node hierarchy and the input field to reach full activation. An input from either one of these sources alone is sufficient only for partial activation, and does not enable the DMF node to generate an output signal. Thus, by priming its subset of the DMF, the winning detector node selects the pattern feature in its receptive field for transmission to the PR system.

Following its allotted time interval of activation, the reset node deactivates the winning detector node for an extended period. This allows the detector nodes to select a new winner, causing a different pattern feature to be presented to the PR system. Continuing in this manner, the BDMD extracts pattern features sequentially and presents them to the PR system for recognition.

Rotation invariance is also facilitated by activation mass detection. The current winning BDMD detector node, acting through the DMF, provides the translation- and scale-invariant image of its receptive field to the DFF. The DFF nodes are connected one-to-one to the RD input nodes. The RD system operates as a *choice* or winner-take-all ART system without learning, having fixed F_2 node templates forming rotated bar patterns. The rotated bar template having maximal ART similarity to the DFF activation pattern determines, through its associated detector node, the RMF subset that will sample the DFF. In a manner similar to that of BDMD detector nodes, the RD detector node primes a fixed-size subset of the RMF. The selected subset of RMF nodes samples the DFF in a manner that effectively inverts the rotation of the associated bar pattern.

4 Getting Smart

In present versions of the architecture, the PR system is an ART system, providing unsupervised learning of feature or object classes. It recognizes familiar, spatially invariant input pattern features presented by the BDMD. The ART categories of the features are accumulated along with their associated scales, translations, and rotations in a working memory.

In the basic architecture, feature selection is performed in a "dumb" mode, where detector nodes compete based upon the mass activity of their receptive fields. The BDMD can be made "smart" in a number of ways. One method is pre-processing of sensor inputs: For example, an input field can represent local inputs from motion detectors, image contour detectors, and other pre-processors operating upon a sensor image pattern. However, a more capable "smart" mode involves feedback to the detector nodes from higher levels of processing in a hierarchical system containing the BDMD. For example, the occurrence of an interesting feature at a large scale could prompt the system to prime detector nodes with smaller receptive fields, yielding higher resolution. This would provide a "zoom-in" capability for interesting features. Also, multiple objects appearing in working memory can be counted by a stack network/ART system combination as described in [2]. The counting system could guide the BDMD in a search for additional occurrences of an interesting object in the input field.

4 Conclusion

The BDMD provides spatial invariance at relatively low cost in terms of neural network computational resources. As described, it can be adapted with additional subsystems and with a smart operating mode, providing flexibility as a pattern recognition front-end system. A simulation of an initial version extracting a feature is shown in Fig. 5.

Acknowledgement

The authors wish to thank Cynthia Actis and Vicki Lane for preparing Figures 1–4.

References

1. Carpenter, G. A. and Grossberg, S. (1987a). "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine", *Computer Vision, Graphics, and Image Processing*, vol. 37,pp. 54-115.

2. Healy, M. J., "On the Semantics of Neural Networks", *Adaptive Neural Systems: The 1990 Technical Report*, T. P. Caudell, Ed. Technical Report BCS-CS-ACS-91-001, Boeing Computer Services, Seattle, WA., 1991.

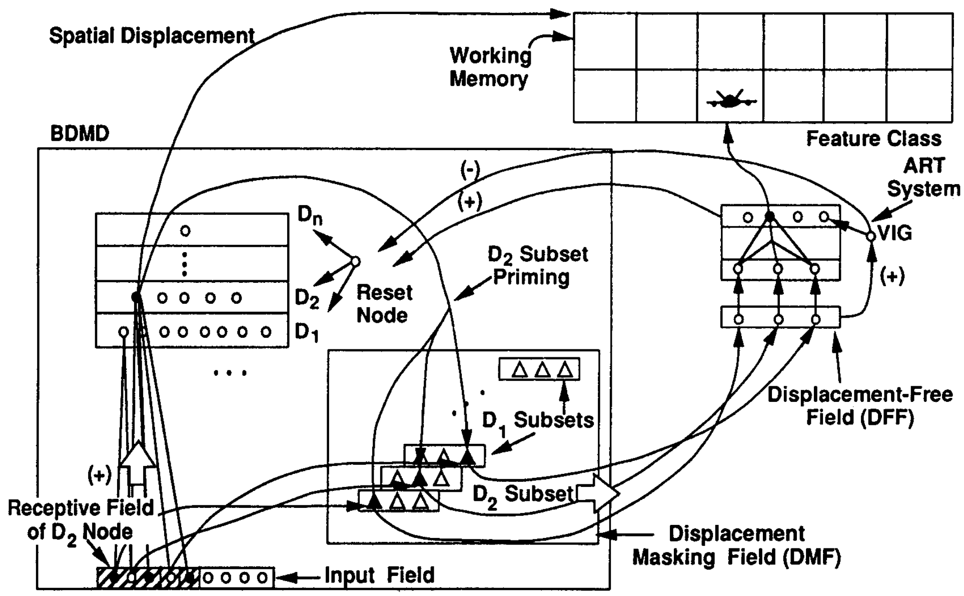

Figure 2. Each detector node primes a DMF subset that samples its receptive field. Low-resolution subsamples of large receptive fields yield fixed feature patterns at the DFF.

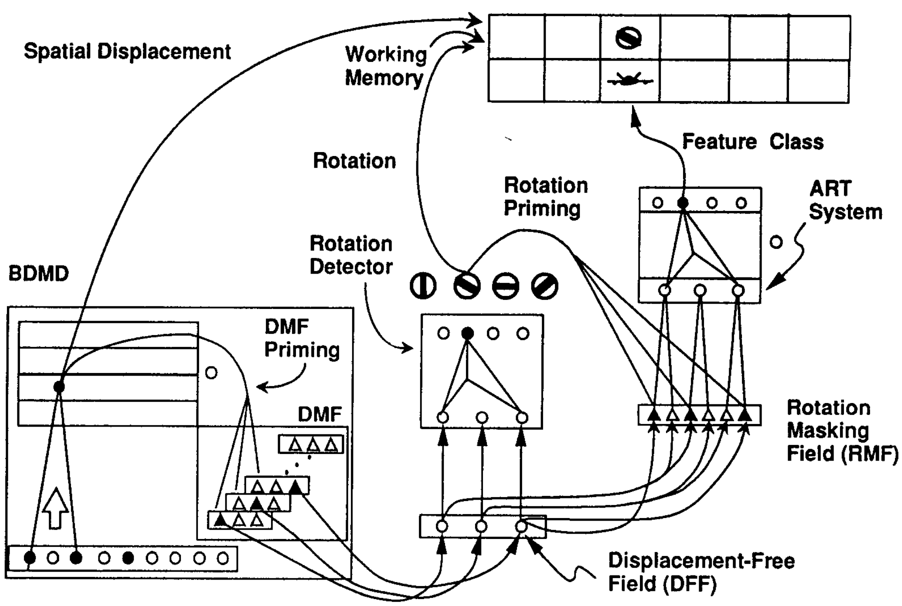

Figure 3. Each RD detector node selects a subset of the RMF that samples the entire DFF. The DFF/RMF connectivity inverts the rotation implicit in the rotation bar template of the detector node.

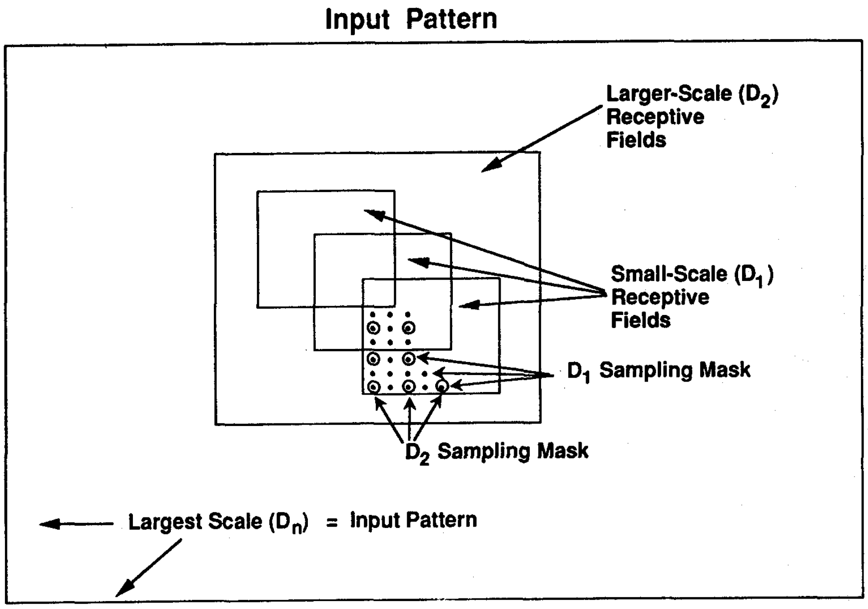

Figure 4. Receptive fields of the detector nodes overlap, and small receptive fields are contained in larger ones.

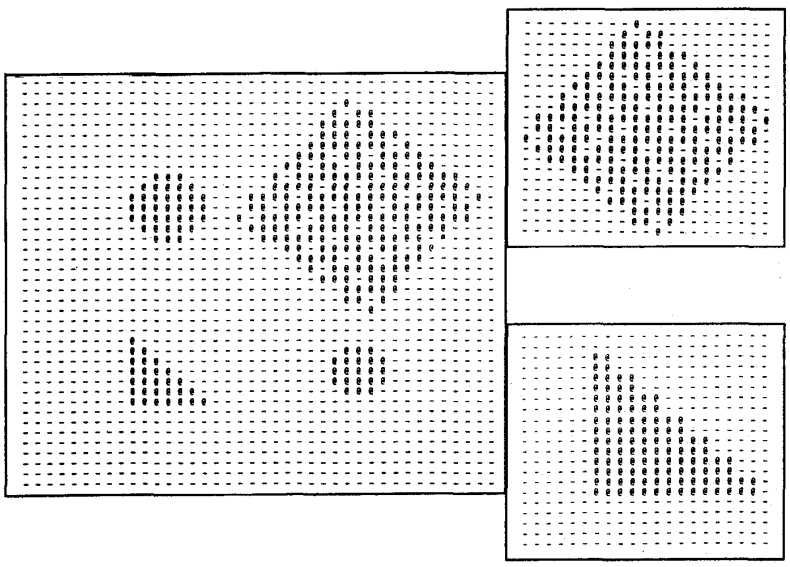

Figure 5. A simulated feature extraction with the BDMD operating upon a binary input pattern.