01 Jan 1991

A Neural Architecture for Unsupervised Learning with Shift, Scale and Rotation Invariance, Efficient Software Simulation Heuristics, and Optoelectronic Implementation

Donald C. Wunsch
Missouri University of Science and Technology, dwunsch@mst.edu

D. S. Newman

T. P. Caudell

R. A. Falk

et. al. For a complete list of authors, see https://scholarsmine.mst.edu/ele_comeng_facwork/1377

Recommended Citation

D. C. Wunsch et al., "A Neural Architecture for Unsupervised Learning with Shift, Scale and Rotation Invariance, Efficient Software Simulation Heuristics, and Optoelectronic Implementation," *Proceedings of the 24th Annual Hawaii International Conference on System Sciences, 1991*, Institute of Electrical and Electronics Engineers (IEEE), Jan 1991.
The definitive version is available at https://doi.org/10.1109/HICSS.1991.183898

A neural architecture for unsupervised learning with shift, scale and rotation invariance, efficient software simulation heuristics, and optoelectronic implementation

Donald C. Wunsch II, David S. Newman, Thomas P. Caudell, David Capps, and R. Aaron Falk

The Boeing Company, P.O. Box 24346, Seattle, WA 98124-0346

Abstract

A simple modification of the adaptive resonance theory (ART) neural network allows shift, scale and rotation invariant learning. We point out that this can be accomplished as a neural architecture by modifying the standard ART with hardwired interconnects that perform a Fourier-Mellin transform, and show how to modify the heuristics for efficient simulation of ART architectures to accomplish the additional innovation. Finally, we discuss the implementation of this in optoelectronic hardware, using a modification of the Van der Lugt optical correlator.

Introduction

The insights presented in this paper come from three sources: the study of adaptive resonance theory (ART)[1-4], the desire to simulate it efficiently in software and the pursuit of optoelectronic implementation. Adaptive resonance theory provides a neural network design to perform unsupervised learning. The binary-input version of this design[2] is referred to as ART1, and we will henceforth assume that we are referring to ART1 in all our discussions. ART1 has been shown to be a type of varying-k-means-clustering algorithm[5] in that it allows patterns to be grouped according to a goodness-of-fit criterion rather than forcing patterns to fit into a preassigned number of categories. The comparison with clustering has led to some efficient algorithms for simplified software simulation[6,7] which we and others have used in previous work.[8,9] Furthermore, we have adapted a classical optical processing technique, called Van der Lugt optical correlation[10,11] to provide an efficient hardware implementation.[12,13] Now we describe how to augment the system with a Fourier-Mellin transform[14,15] so as to provide the desired invariances. The combination of ART and the Fourier-Mellin transform has been reported previously.[16] Our main contribution is to point out that this can be efficiently incorporated into the hardware implementation, since the hardware we are using is particularly amenable to this technique[13].

Unsupervised learning with ART

Adaptive resonance is a very well-known neural network theory.[1-4] The reason for its fame in the neural network community is that it uses very simple elements to perform learning without a teacher. In other words, there is no training signal presented along with each input pattern that allows the network to learn a proper output. The network must learn the proper response without assistance. However, this is not the only useful feature of ART architectures. They also have the property that they regulate their own learning. There is no signal needed to tell an ART network to switch from a "learning mode" to a "performance mode". Finally, the network is stable, yet always ready to learn something new. It can learn a set of patterns, then get dealt some novel patterns, deal with those in an appropriate manner, and yet retain a reasonable categorization of the old memories. No other neural net theory can boast all of these capabilities.

How does an Adaptive Resonance Unit do all this? The key is that the pattern classification takes place in a feedback loop, and that learning does not appreciably set in until resonance occurs. If resonance does not occur, there is a mechanism called reset that allows a search for a better pattern match, removes all previously considered classifications, and suspends learning until the right answer is found. This is clarified by examining figure 1.

In figure 1 we see the ART unit displayed in several separate layers. These are: R, the recognition layer; C, the comparison layer; I, the input layer; V, the vigilance layer; and Re, the reset layer. This grouping of layers is taken from a paper by Ryan et. al.[6], and while it does not follow Carpenter and Grossberg's description exactly, it is topologically and functionally equivalent. Going left-to-right we see

0073-1129/91/0000/0298$01.00 © 1991 IEEE

the ART unit in action. First the input is merely registered at the comparison layer and fed up to the recognition layer (1a.) In the second frame, we see that the recognition layer has a winner-take-all property, so that the node corresponding to the best match wins and is the initial best guess (1b.) This guess is not taken as the final word, though, as we see at the bottom left. It is instead tested by playing back the previously learned template associated with the winning node onto the comparison layer. This is compared with the pattern still on the input layer by competing signals sent to the vigilance layer (1c.) In the final frame we see an example of what happens when the match is not good enough. The vigilance layer is now able to activate the reset layer. The reset layer has the property that it only suppresses nodes at the output that have been recently active, and has no effect on the rest (1d.) In this case, only the prior winner has been affected. Now with that node removed, the network will reclassify the pattern and continue to do so until it has found a good match.

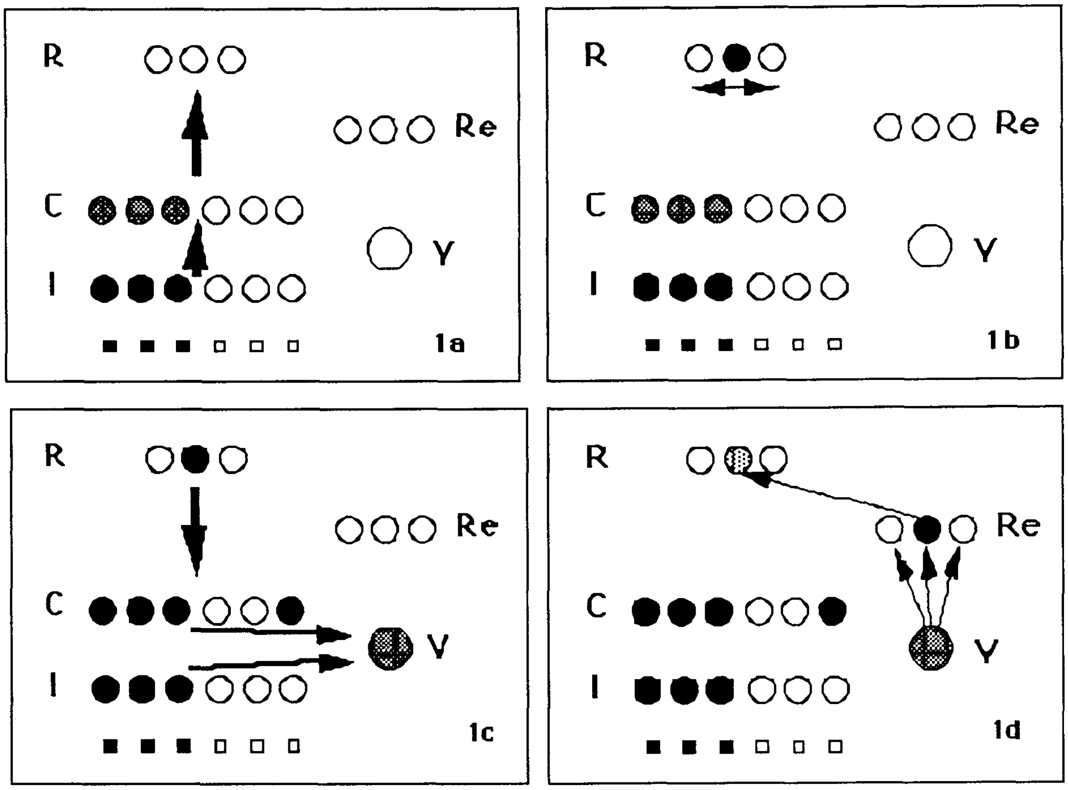


Figure 1. The learning cycles of an ART unit.

The decisions that the unit performs at the various points in this cycle have a simple mathematical characterization (in contrast with the dynamical equations for the system, which are highly coupled and nonlinear.) There are fortunately many theorems about the system, and some of them show some simple decision rules for various parts of the system.[8] For example, the recognition layer simply chooses the node that has the greatest normalized inner product between the input pattern and the template pattern. Similarly, the vigilance check is determined by taking the same inner product, with a different normalizing factor. This procedure is specified more rigorously below.

Consider a new n-element input vector to be called P. (All patterns are referred to as vectors here, even though they will be two-dimensional patterns in most experiments, and could easily take on three or more dimensions. It does not change any of the following analysis to simply consider the patterns as vectors.) Now consider how each of these vectors will be classified. We wish to assign the vector to a category, say category 1, category 2, etc. Each category will have some template, or prototype vector, associated with it. These have also been referred to as library elements in the literature. They are the patterns that the unit "knows". The unit will compare the input to these patterns to decide how to classify it. We will index these patterns and refer to them as the T_i, where i is the index number. The parameter ρ is the vigilance threshold of the ART unit, and β is a small bias. With these definitions in mind, the full ART1 algorithm is given in the flowchart below.[12]

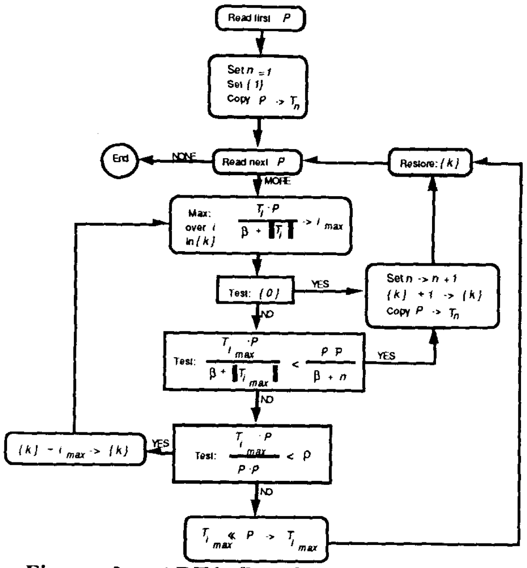
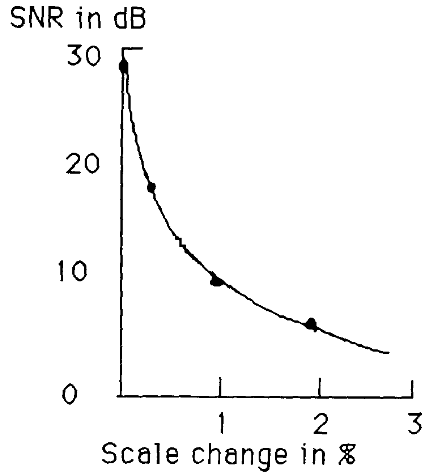


Figure 2. ART1 flowchart

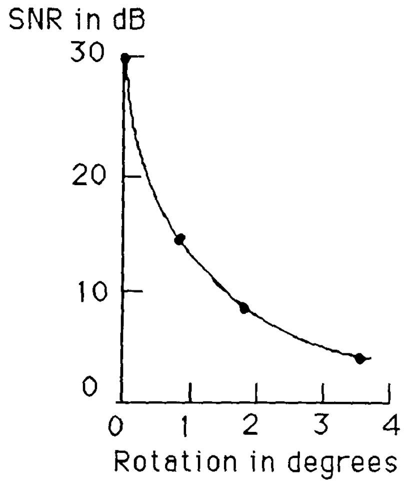
The Fourier-Mellin transform

In 1964, a powerful technique in optical computing was introduced. It allowed the all-optical recognition of a pattern, even in the presence of considerable noise. This technique is known as Vander Lugt correlation. Despite its power, it still has not made the optical pattern recognizer an off-the-shelf device. One reason for this is that the device has the following problem: a small shift, rotation, or change in scale results in a large degradation of the system's performance. An illustration of the latter two problems (which can be regarded as the more serious ones) is shown in figure 3 (adapted from Casasent and Psaltis.[14, 15]) From this figure we see the marked decrease in signal-to-noise-ratio (SNR) for a few degrees rotation or a few percent scale change.

A promising solution to this problem was proposed by Casasent and Psaltis.[14, 15] This method uses optical transformations which provide invariance to position, rotation, and scale changes. The first step in this procedure is to form the magnitude of the Fourier transforms of the input object (to be recognized) and of the "ideal" object (to be stored as a



3a. Degradation of SNR with scale[15]



3b. Degradation of SNR with rotation[15]

template.) This step can be realized by a Fourier transform lens. The next step is to convert these functions to polar coordinates (i.e. $F_1(r,\theta)$ and $F_2(r,\theta)$.) Then these functions are scaled logrithmically in r.

These steps (which are really combined into only one step) can be performed by computer generated holograms or spatial light modulators. The final step is to perform the Fourier transform of the scaled functions. The combination of log-polar scaling and Fourier transform is equivalent to a Mellin transform. With the Fourier transform at the front, the entire process is referred to as a Fourier-Mellin transform and the output patterns are invariant to position, scale and rotation changes of the input patterns.

The neural architecture

The design of a hardwired neural network to perform a discrete Fourier transform (DFT) is a straightforward exercise, since a DFT is simply an appropriate summation of weighted inputs.[17] Furthermore, a Mellin transform can also be designed as a hardwired remapping of patterns.[14,15] Therefore, the Fourier-Mellin transform can be accomplished as a series of hardwired neural networks. The entire neural net architecture for position, scale, and rotation invariant unsupervised learning can therefore be considered as a linking of neural net components.

Efficient software simulation

The ART1 algorithm shown in the flowchart (figure 1) is based directly on the theorems in Carpenter and Grossberg[2]. It has been verified by Carpenter[18], and the Boeing simulation code implementation has been tested against the published examples in[2], yielding identical results.

The software investigations of ART1 performance on a variety of data which are aimed at delineating the tasks it performs most effectively will also provide a database for testing hardware performance. One database which is currently under investigation is generated to satisfy the necessary and sufficient conditions for perfect ART1 learning of the theorem proved in Newman and Caudell[17]. It consists of a series of hierarchically clustered classes of patterns which will be learned perfectly on the first pass through the data. The classification becomes finer and finer as the data is presented to ART1 units of increasing vigilance level. The alphabet example[2] is also investigated in greater detail using the statistical techniques suggested in Newman and Caudell[17].

Hardware implementation

The basic configuration of the electro-optical ART unit is a type of Van der Lugt correlator[5], given in figure 4. Plane P1, in the center, is the input plane. A spatial light modulator is configured as a binary phase-only filter, and contains the two-dimensional Fourier transform of the input pattern. Plane P2 is the template plane. This is also a spatial

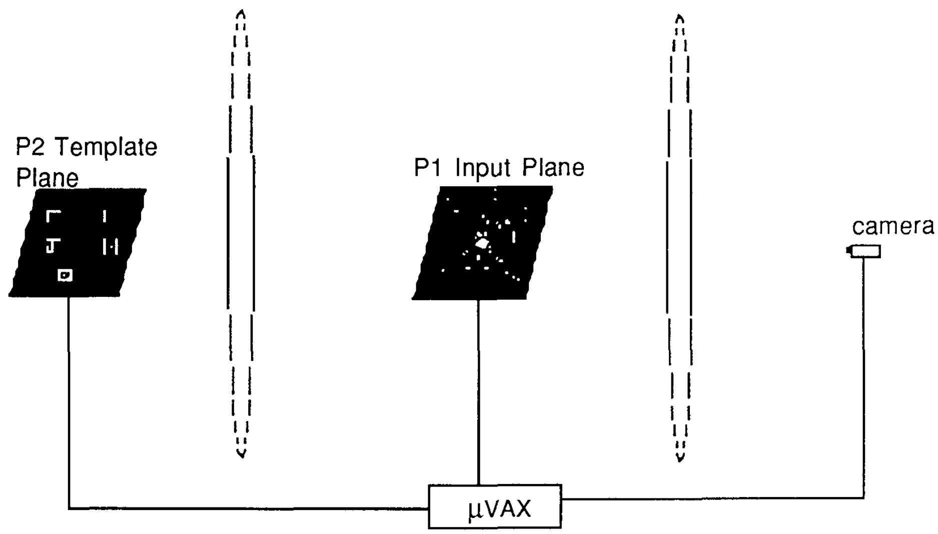


Figure 4. The basic system.

light modulator, but it is an amplitude modulator. It contains multiple templates simultaneously. The lenses are chosen such that the paraxial approximation applies, and the plane P1 will receive the Fourier transforms of each template, with the zero frequency spot in approximately the same location.

The output plane is P3, where we have a charge-coupled-device (CCD) camera. This plane will have correlation peaks corresponding to templates that closely match the input image, as shown in figure 5. It is possible to sample this plane so that sample points will correspond to the value of $|p \cdot t|$, where p is the pattern and t is the template, both expressed as one-dimensional vectors. This is a key part of what an adaptive resonance unit calculates, as we saw in the flowchart of figure 1. Other key information needed is $|p \cdot p|$ and $|t \cdot t|$. These can easily be calculated, either optically or electronically. For example, the $|p \cdot p|$ term can be calculated by including a copy of p on the template plane, and measuring its corresponding output at the camera. The $|t \cdot t|$ term is not recalculated as frequently and may be easier to do electronically, although it could be easily be done optically with this system by time-multiplexing the calculations. Also, the output is calibrated electronically to compensate for image degradation due to intensity variations across the template field.

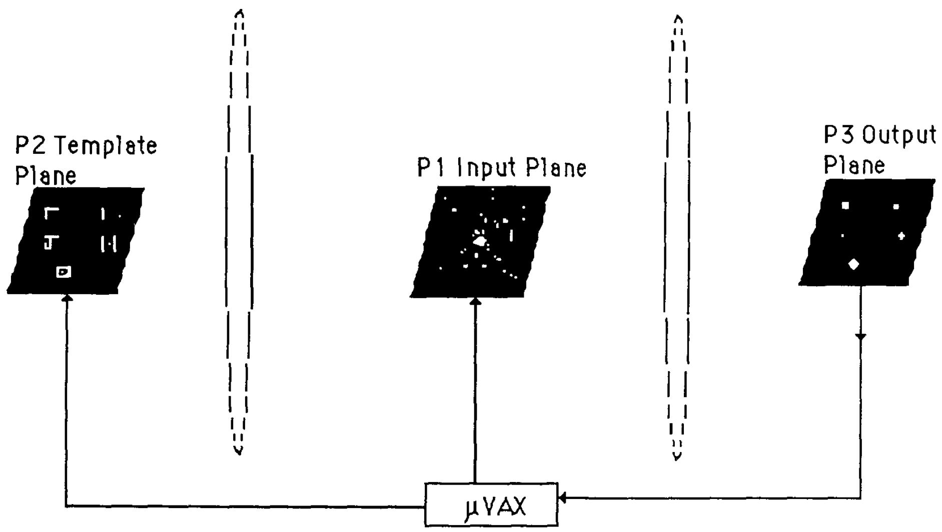


Figure 5. Multiple matched filter correlations.

This is a proven experimental setup for performing other kinds of optical computing operations.[5,7,8] All electronic calculations, and control of the spatial light modulators, is done by a DEC microvax. The unit is also used to process the CCD camera output via a frame grabber. The control code is written in FORTRAN and also calls some special control macros for the spatial light modulators. A photograph of the experimental setup is shown in figure 6.

This system is especially attractive for the problem at hand: shift, scale, and rotation-invariant correlation. The reason for this is that we are already using a correlator that forms the backbone of a generalized optical transform system, that is easily capable of performing the Fourier-Mellin transform as described by Casasent and Psaltis[14,15]. Modification of the correlator can take the form of inserting a steering phase element in the Fourier plane. However, it can be implemented even more simply. The spatial light modulator that is already in that position can simultaneously act as a steering phase element and input element when programmed correctly. Techniques for calculating the generalized phase element for a desired optical transform are given by Saleh and Freeman[20].

It is appropriate to caution the reader that the shift invariance provided by this device is limited to small shifts. This is because a shift in the input causes a corresponding shift in the output correlation peak. This is normally of no consequence, since the maximum peak can be noted wherever it occurs. However, when the correlator is configured to process multiple patterns simultaneously, as in this device, the shifts must be small enough that the correlation peak does not move into a region associated with a neighboring library element. The rotation and scale invariant properties are not limited by this property--only the shift invariance is affected.

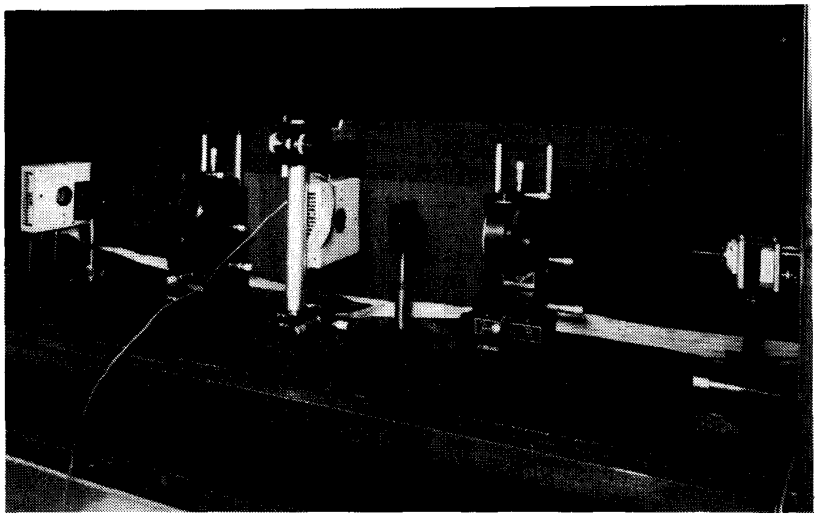


Figure 6. The experimental apparatus.

Conclusions

We have described a system that is capable of position, scale, and rotation invariant unsupervised learning. This is accomplished by coupling an adaptive resonance unit with a Fourier-Mellin transform. Furthermore, we show how this can be described as a completely neural system, and discuss simulation and hardware implementation issues.

References

1. Grossberg, Stephen, "Competitive learning: from interactive activation to adaptive resonance," Cognitive Science, 1987, No. 11, pp. 23-63, Cognitive Science Society

2. Carpenter, Gail A., and Stephen Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," Computer Vision, Graphics, and Image Processing, 1987, No. 37, pp. 54-115, Academic Press, Inc.

3. Hestenes, David, "How the brain works: the next great scientific revolution," in Neural Networks: An Introduction, Donald Wunsch II, Ed., Pergamon Press, submitted

4. Grossberg, Stephen (editor,) Studies of Mind and Brain, D. Reidel: Dordrecht, Holland, 1982

5. Burke, Laura Ignazzio-, submitted to Neural Networks

6. Ryan, T. W., and C. L. Winter, "Variations on adaptive resonance," SAIC

7. Moore, Barbara, "ART1 and pattern clustering," Proceedings of the 1988 Connectionist Summer School at Carnegie-Mellon University, Touretzky and Hinton, Eds., Morgan Kaufmann Inc., 1989

8. Healy, Michael J., "The elements of adaptive neural expert systems," SPIE Vol. 1095, Applications of Artificial Intelligence VII, pp. 830-837, SPIE, 1989

9. Thomas P. Caudell, Donald C. Wunsch II, Richard A. Escobedo and Stanley K. Tazuma, "Visual clustering of aircraft part designs with adaptive resonance neural networks", submitted to Optics Letters

10. Van der Lugt, A. B., "Signal detection by complex spatial filtering," IEEE Transactions on Information Theory, IT-10, No. 2, IEEE, 1964

11. Goodman, Joseph W., Introduction to Fourier Optics, pp. 141-184, McGraw-Hill, San Francisco, 1968

12. Wunsch II, Donald C., Thomas P. Caudell, R. Aaron Falk, and C. David Capps, "An optoelectronic learning machine," OSA Annual meeting 1990, Boston, November 1990

13. Wunsch II, Donald C., and Thomas P. Caudell, "An optoelectronic learning machine," Proc. Northcon, October 9-11, 1990, Electronic Conventions Management, 1990

14. Psaltis, Demetri, and David Casasent, "Position, Rotation, and Scale Invariant Optical Correlation", Applied Optics, Vol. 15, No. 7, July 1976

15. Psaltis, Demetri, and David Casasent, "New Optical Transforms for Pattern Recognition", Proc. IEEE, Vol. 65, No. 1, January 1977

16. Carpenter, Gail A., and Stephen Grossberg, "Invariant pattern recognition and recall by an attentive self-organizing ART architecture in a nonstationary world," Proceedings of the IEEE First International Conference on Neural Networks, San Diego, California, June 21-24, 1987, IEEE Catalog Number 87TH0191-7

17. Culhane, Andrew D., Martin C. Peckerar, and C. R. K. Marrian, "A neural net approach to discrete Hartley and Fourier Transforms," IEEE Transactions on Circuits and Systems, Vol. 36, No. 5, May 1989, pp. 695-703

18. Personal communication.

19. Newman, David S. and Thomas P. Caudell, "Characterization of some learning properties of ART machines," Neural Networks for Automatic Target Recognition, Wan Institute of Boston University, May 1990

20. Saleh, Bahaa E. A., and Mark O. Freeman, "Optical Transformations," in Optical Signal Processing, Joseph L. Horner, editor, Academic Press, New York, 1987, pp. 281-333