VIGILANTE: An Advanced Sensing/Processing Testbed for ATR Applications

Suraphol Udomkesmalee, Anil Thakoor, Curtis Padgett, Taher Daud, and Wai-Chi Fang
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California 91109

and

Steven C. Suddarth
Ballistic Missile Defense Organization
Washington, DC 20301

ABSTRACT

VIGILANTE consists of two major components:

1) the **V**iewing **I**mager/**G**imballed Instrumentation Laboratory (VIGIL)—advanced infrared, visible, and ultraviolet sensors with appropriate optics and camera electronics

2) the Analog **N**eural **T**hree-dimensional processing Experiment (ANTE)—a massively parallel, neural network-based, high-speed processor,

The powerful combination of VIGIL and ANTE will provide real-time target recognition/tracking capability suitable for Ballistic Missile Defense Organization (BMDO) applications as well as a host of other civil and military uses.

In this paper, we describe VIGILANTE and its application to typical automatic target recognition (ATR) applications (e.g., aircraft/missile detection, classification, and tracking), this includes a discussion of the VIGILANTE architecture with its unusual blend of experimental 3D electronic circuitry, custom design and commercial parallel processing components, as well as VIGILANTE's ability to handle a wide variety of algorithms which make extensive use of convolutions and neural networks. Our paper also presents examples and numerical results.

Keywords: automatic target recognition, neural processor, multispectral sensors, system architecture

1. INTRODUCTION

In-flight, small, "brilliant" systems capable of autonomously acquiring and identifying hostile targets (e.g., cruise missiles, missile launchers, and other types) are an essential component of **BMDO's** planned defensive mechanisms. Such an Automatic Target Recognition **(ATR)**[1] capability could greatly enhance the probability of mission success for interceptors (for cruise and ballistic missiles), surveillance platforms (for missile launchers), and ground-based fire control, The ability to autonomously and **efficiently** seek and destroy such threats is particularly important as we consider their **widescale** proliferation and technical advancement. VIGILANTE will' provide a flexible/portable, low-cost testbed needed to evolve intelligent seeker systems (consisting of sensors, optics, and processing) for BMDO applications.

Problems associated with ATR have given conundrums to researchers for decades. Despite recent progress in microprocessor technology, no deployable ATR systems capable of providing reliable target recognition and tracking in real time exist to-date for autonomous defense weapon systems, Parallel computer systems capable of giga-operations/s, such as Adaptive Solutions' CNAPS array **processors,**[2-3] could perform convolution with a small kernel (3x3 or 8x8) in real time, but achieving general object recognition from video in real time with a reasonable template size (32x32 or larger) is still beyond to&y's computer/processor technologies. Optical **correlators** designed for distortion-invariant pattern recognition are an attractive alternative for achieving massively parallel processing with photons, but they have yet to overcome many performance and system issues, such as signal-to-noise, discrimination ability, **programmability** and limitations of available spatial light modulators, post-processing of correlation outputs, and packaging, that prevent the realization of a flexible, robust optical processor for ATR.

Sensor technologies needed to provide capable "eyes" for ATR have made, to some extent, excellent progress in providing multiwavelength target data and engendering the field of sensor fusion, both of which are needed for handling multisensory **environments.**[5] A typical surveillance platform like the Midcourse Space Experiment **(MSX)**[6] provides multiwavelength **phenomenology** measurements for development and assessment of target discrimination algorithms.

VIGILANTE will, for the first time ever, provide a complete multisensory and processing system in a small package that is suitable for ATR. An integrated testbed will incorporate an advanced artificial neural network processor and several new sensors covering the spectral range from ultraviolet **(UV)** through infrared **(IR).** Periodically during this development, the instrument systems will be flown on an airborne platform to obtain data for developing and evaluating system effectiveness, The flight tests will provide realistic data sets (targets and environments) not achievable in the laboratory. The new, lightweight sensors are the Quantum Well **Infrared** Photodetector **(QWIP),** the Active Pixel Sensor **(APS),** and the delta-doped ultraviolet charge-coupled device **(UV CCD).** These three sensors cover the wavelength ranges 8 to 9, 0,5 to 0.9 and 0.3 to 0.7 **μm,** respectively. On the image data-processing side, **the neural** network advances have led to chip implementations that are currently being assembled into innovative, three-dimensional architectures capable of tera-operations/s and running 64 image-based convolutions (with 64x64 kernel size) in real time.

VIGILANTE will pave the way for unique onboard, real-time processing of sensor images for autonomous interceptors and general-surveillance systems. Real-time target recognition will be demonstrated through a series of ground/airborne experiments using real target images.

2. SYSTEM DESCRIPTION

VIGILANTE consists of the Viewing Imager/Gimballed Instrumentation Laboratory (VIGIL) and Analog Neural Dee-dimensional processing Experiment (ANTE). VIGIL is an airborne telescope serving the dual functions of data acquisition for target recognition experiments and testing of novel active and passive focal plane imagers. The telescope will consist of a self-contained 15-cm Cassegrain unit, a gimbaled mirror, and channels for multiband sensors. A schematic diagram of the VIGILANTE system is shown in Figure 1.

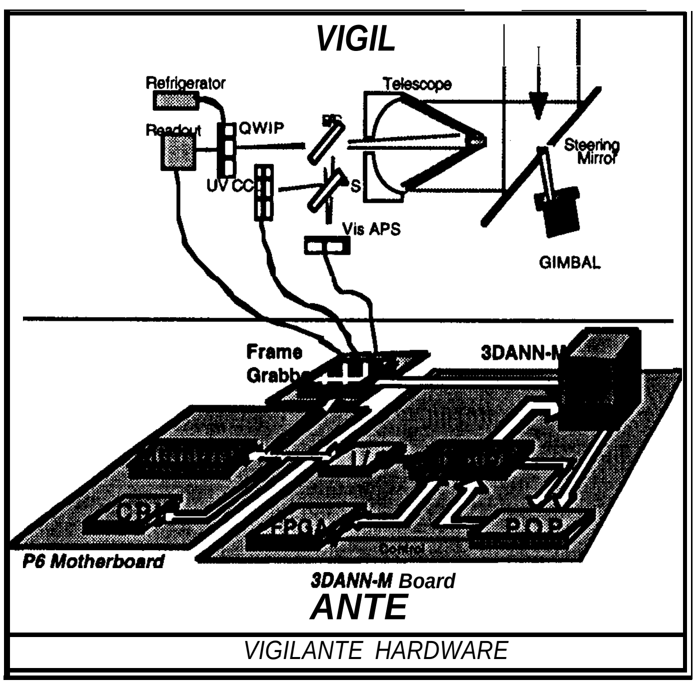


***Figure 1.* A schematic diagram of the VIGILANTE system. VIGIL is an integrated optical system that splits/transmits the incoming light (steered by a gimbaled mirror) detected by the respective IR/visible/UV sensors. ANTE is the processing system that selects each sensor channel for processing that is done by a commercial frame buffer and host processor and carries out real-time ATR by means of specialized, analog neural networks (3DANN-M) and a point operation processor (POP).**

ANTE is a prototype ground and airborne image-processing/target-recognition computer architecture based upon technology developed under the ongoing 3-dimensional artificial neural network (3DANN) program. 3DANN is a sugar-cube-sized, low-power neuroprocessor with its IC stack mated to an IR sensor array.[7-8] ANTE uses a modified version of the 3DANN referred to as 3DANN-M, which allows VIGILANTE to accept data from any sensor of arbitrary size and format, More importantly, the 3DANN-M cube can be used for general image convolutions.

3. SENSORS

VIGILANTE will assess capabilities important to BMDO for the following advanced passive sensors:

1) The Quantum Well Infrared Photodetector (QWIP)[9] is a 256x256-pixel, real-time (30-to 120-Hz frame rate) IR sensor array, QWIP is a major advance over the state-of-the-art HgCdTe sensors. The sensor has an 8- to 9-pm central wavelength detection capability with a 1 µm full width at half maximum, and the design uses a random reflector on each pixel to maximize light trapping. The main advantages of QWIP sensors, which are based on GaAs, over HgCdTe sensors, are

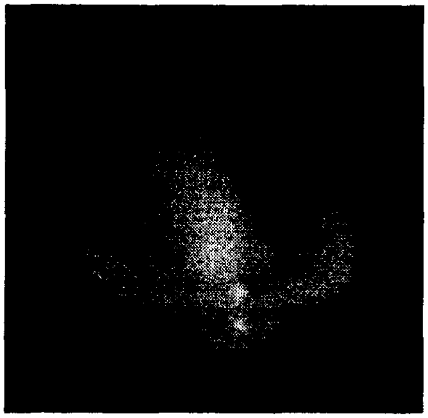
- Excellent stability and easier processing from using GaAs
- Exceptionally good array uniformity (in excess of 99%)
- Higher array yields
- Easy scalability to arrays of larger sizes
- Sensitivity/tunability over a wide wavelength range (~6 to 19 µm)
- Possible fabrication of dual-band arrays

2) The Active Pixel Sensor (APS)[10] visible camera system is based on second-generation solid-state image sensor technology. It retains nearly all the performance of a charge-coupled device (CCD), with the added enhancements of ultralow power (100 times less than is used by a CCD system), 0,5- to 0.9-pm sensitivity, faster readout, radiation hardness, and arrays in larger formats. The demonstrated APS array size is 256x256, with a pixel rate of 1 mega-pixel/s (15 frames per second). The CMOS chip has standard 5-volt or, alternately, 3.3-volt operation with a power requirement of 3 to 30 milliwatts.

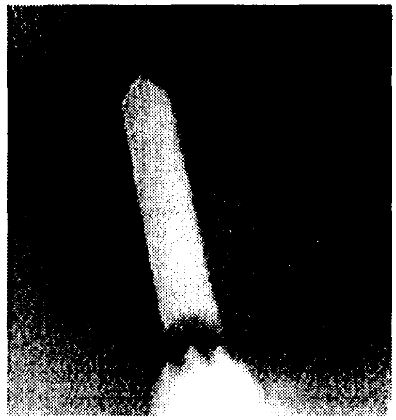
3) The UV CCD Camera[11] is based on delta-doped CCD technology. It consists of a doped-silicon layer 2.5 nanometers thick that is grown on the backside of a thinned CCD using Molecular Beam Epitaxy (MBE). Delta-doping enhances detection of electrons generated by UV photons to almost 100% efficiency. An antireflection coating is added to the sensor array for enhanced sensitivity for 0.3 to 0,7 µm. The frame transfer active area is 256x5 12 pixels with a 30 frame/s effective speed. A suitable electronics package is added to the camera for its operation.

The optical system for the detectors described above consists of the telescope, the beam splitters, and the supporting structure (see Figure 1). The telescope is a standard 15-cm Ritchey-Chretien design. It receives its input from a gimbaled mirror that reflects light to the telescope primary first and then to the secondary mirror before directing it to the beam splitters. The telescope has a 20-milliradian (1 .2-degree) field of view. The primary beam splitter is of multilayer construction, is made from common coating materials, and is designed to operate with a 45° angle of incidence. It is designed to reflect the UV/visible wavelengths and transmit the IR. Reflectance in the UV is 71 % between 0.33 and 0.38 µm and is 95% in the visible (0.5 to 0.7 µm). Transmission in the IR is 90% for 3 to 3.7, 65% for 4 to 5, and 80% for 7 to 11 µm. The secondary beam splitter is similarly designed to reflect UV and transmit visible wavelengths. The focal length for the entire f/1 0.7 system is 160 cm.

VIGILANTE's sensors can be queued to assist in the ATR **functions** of detection, classification, and **precision-** tracking. For example, the UV wavelengths (0.3 to 0.7 **μm)** can be used for detection of plumes from BMDO targets of **interest**[12]**; IR** (8 to 9 **μm)** is suitable for cold-body sensing and permits classification of these targets (see Figure 2), and the visible wavelengths (0.5 to 0,9 pm) can be used for close-up tracking to provide aim-point selection for the end-game scenario, Eventually, the VIGILANTE sensors may be used for simultaneous **fusion** of the data from all wavelengths.



(a)



(b)

Figure 2. Hot-plume/cold-body flux difference: sensing of the same target at (a) 3- to 5-μm infrared, with hot-plume/cold-body flux ratio of 25,000, and (b) 8- to 9-μm infrared, with hot-plume/cold-body flux ratio of 114 (QWIP image), Owing to the significantly smaller flux ratio, the QWIP sensor is able to provide images of the cold-body in the presence of the hot-plume, which cannot be done by the 3- to S-pm sensor.

4. THE ANALOG NEURAL NETWORK

The main contribution in VIGILANTE is the ANTE processor, a unique combination of an experimental 3D **neural** IC, custom **circuitry** and an off-the-shelf parallel processor, The system was particularly designed to recognize shapes in resolved images at extremely high speeds (on the order of **tera-OPS).**

The heart of the ANTE processor is the 3DANN-M neural **"sugarcube"** chip stack. A neuron can be simply modeled by a nonlinear threshold unit fed by a linear combination of inputs. 13 3DANN uses a neural circuit design based on Multiplying Digital-to-Analog Converter **(MDAC)** technology; each circuit is digitally programmable, has 8-bit resolution digital weight storage, and is an analog multiplier with a **voltage-input/current-output** configuration, [14] This innovative design reduces the transistor count by half for the same bit resolution., This increases speed, decreases chip size, and increases from a 7-bit to an 8-bit synapse cell for a 64x64 synapse array. The **block** diagram of 3DANN-M is shown in Figure 3. In 3DANN-M, 64 complete **neural** inner products, each with a 4096 (i.e., 64x64) input array can be accomplished in **250** nanoseconds (i.e., 10^{12} multiply and add operations in 1 second, or 1000 **frames** per second). The **3DANN-M** circuitry is designed to operate at 90 K temperature and has a low power consumption of approximately only 2.5 watts. The threshold is excluded to enable general inner-product and convolution operation on **3DANN-M.**

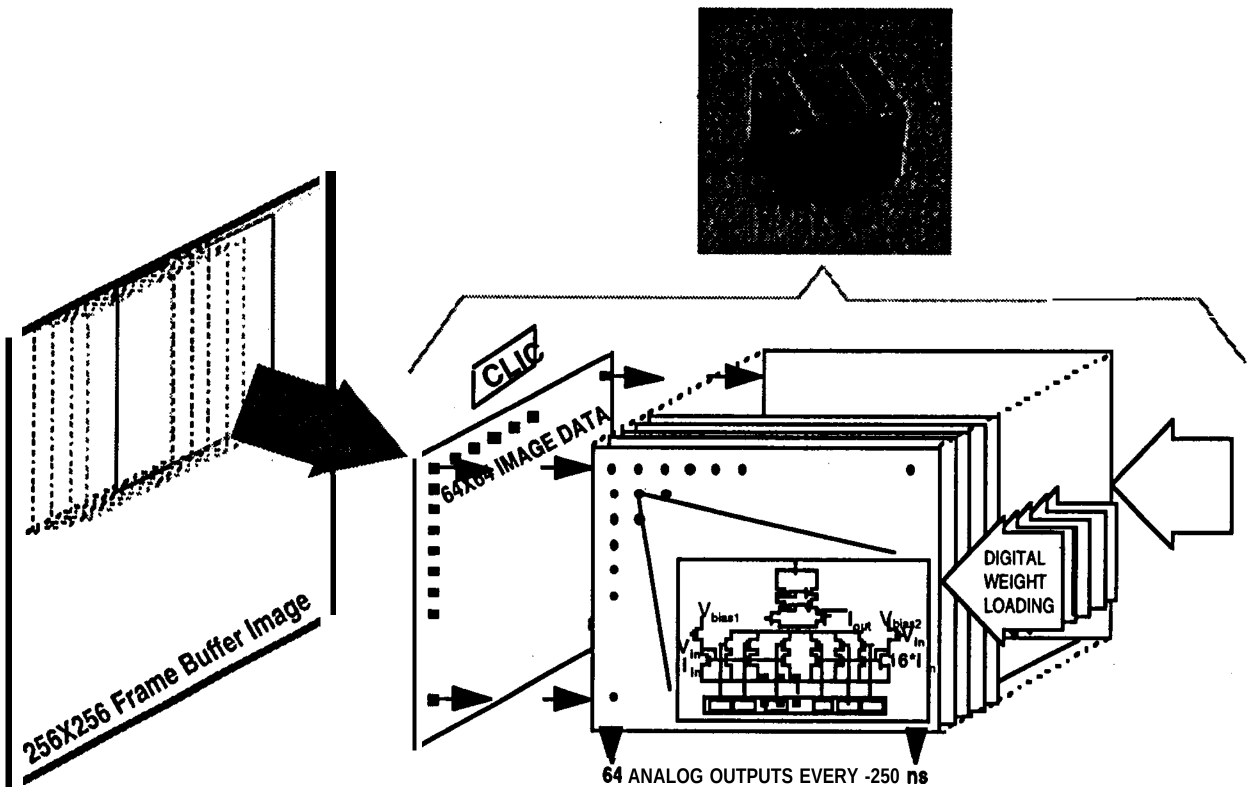


Figure 3. The 3DANN-M network consists of 64 layers of a 64x64 synapse array based on an 8-bit MDAC. It incorporates a special-purpose image-write device called CLIC that is bump-bonded to the synapse array stack. 3DANN-M can be realized (as pictured) in a 10-gm, 3-cm^3 package, with power consumption of 4.5 W. CLIC is specially designed for rastering of a 64x64 window of a larger image in the frame buffer and is synchronized with 3DANN-M's 250-ns inner-product operations.

Using the Column Loading Input Chip **(CLIC)** device, a 64x64 block of image data can be **rastered** and **fed** through for processing in the 64 layers of synapse circuit arrays. A schematic diagram of CLIC is shown in Figure 4. It shows a 64x64 static random access memory **(SRAM) array** consisting of 8-bit shift registers **(8-SHR);** these store the input signals as 8-bit digitally stored weights and shit? the stored signals, column by column or row by row, to the right, down, or up as required, with **the** shift signals SRC, SDC, or SUC, respectively. [15] The new gaps are **filled** one row (or column) at a time from the external **frame** buffer. The stored weights are converted by the Digital-to-Analog Converters **(DAC's)** to analog voltage signals. (Shown next to the solid squares in Figure 4). The 64x64 solid squares **signify** the positions of the iridium bumps by which the voltage signals will be transferred in parallel to their respective inputs of the 3DANN-M stack.

Incorporating a synapse circuit that is modeled after a biological neural network, makes the sugar-cube-sized **3DANN-M** an extremely powerful image-processing engine capable of carrying out in parallel 64 convolutions of the form

$$c_i(x,y) = f(x,y) \otimes g_i(x,y); \; i = 1, 2, \ldots 64; \qquad (1)$$

where **f is** the input image, g_i is the filter mask, and c_i is the output image,

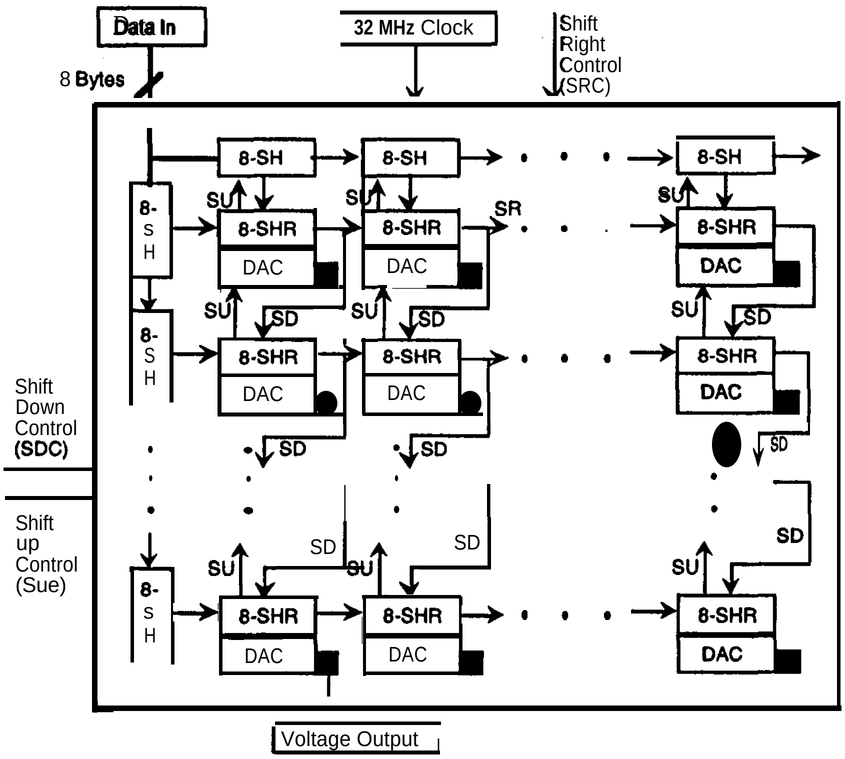


Figure 4. Block diagram of the Column Loading Input Chip (CLIC) showing the SRAM 8-bit shift registers (8-SHRs), the DACs, and the iridium bumps that mate it with the 3-D stack of synapse array chips. The initial column (or row) storage of image data will be done in the respective 8-bit shift registers (8-SH) during the time the signal processing of the previously loaded inputs is proceeding** in **the 3DANN-M module. The actual column-wise or row-wise shifts will be achieved within a maximum time of 33 ns, because of the 32-MHz clock rate and 8-bit bandwidth. Thus, the high speed of 3DANN-M needed for data processing within 250 ns will be maintained.

5. ANTE PROCESSOR ARCHITECTURE

The general ATR process flow is depicted in Figure 5. The 3DANN-M network produces 64 inner-products (each with two 4096-element vectors) every 250 ns, so the frame buffer is used to hold the image and feed a new column or row of a 64x64 subwindow to CLIC every 250 ns (thus accomplishing 64 convolutions of a 256x256 image with 64x64 masks in 16 ins). The 64 analog values generated by 3DANN-M every 250 ns are converted to 8-bit digita! values and passed along to the associated feedback memory and Point Operation Processor ('FOP).

Currently, the feedback memory and POP are implemented in **VIGILANTE** with a commercial product—four of Adaptive Solutions' CNAPS array processor boards (each board containing 128 SIMD processors and 32 megabytes of **memory)—providing** flexibility in programming different point processing operations. In later stages of the project, a custom VLSI implementation of POP may be designed and fabricated. POP takes the output from the 3DANN-M and performs the desired target recognition.hracking **functions.** Command and control of VIGILANTE operations (e.g., **detection/classification/tracking** mode command, loading of templates, point operation functions, data recording, etc.) are done though the P6 motherboard (shown as the processor/memory block in Figure 5).

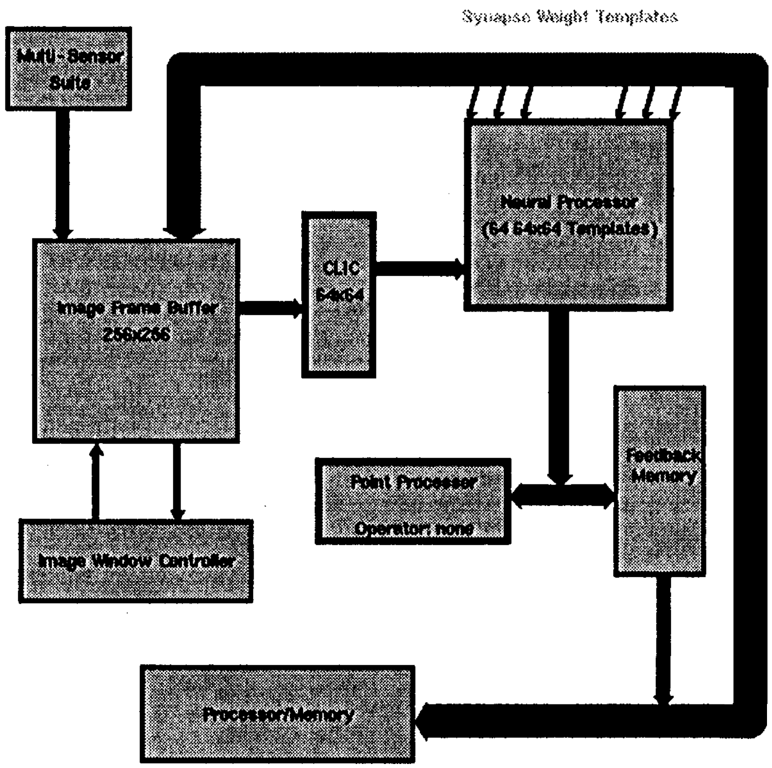


Figure 5. The VIGILANTE processing architecture that orchestrates the data flow from sensor through neural processor also serves as the basis for developing methodologies for A TR applications.

6. ALGORITHMS AND SIMULATION

To **efficiently** recognize objects of arbitrary size and orientation, a hierarchical neural network approach based on **eigenvectors** is employed (see Figure 6). Using 3DANN-M as the dedicated synapse weight multiplier hardware, 64 **eigenvector** templates representing the principle axes of a collection of multidimensional data points (i.e., object images of various configurations) are employed at a time. 16-18 Since each **data** point (image) is a 4096-element vector, finding a set of 4096 **orthonormal eigenvectors** is possible (64 of which can reside on **3DANN-M).**

Selecting the most significant 64 **eigenvectors** constructed **from** principle component analysis of target imagery reduces the **dimensionality** of image sets, yet still retain much of the information relevant for classification.

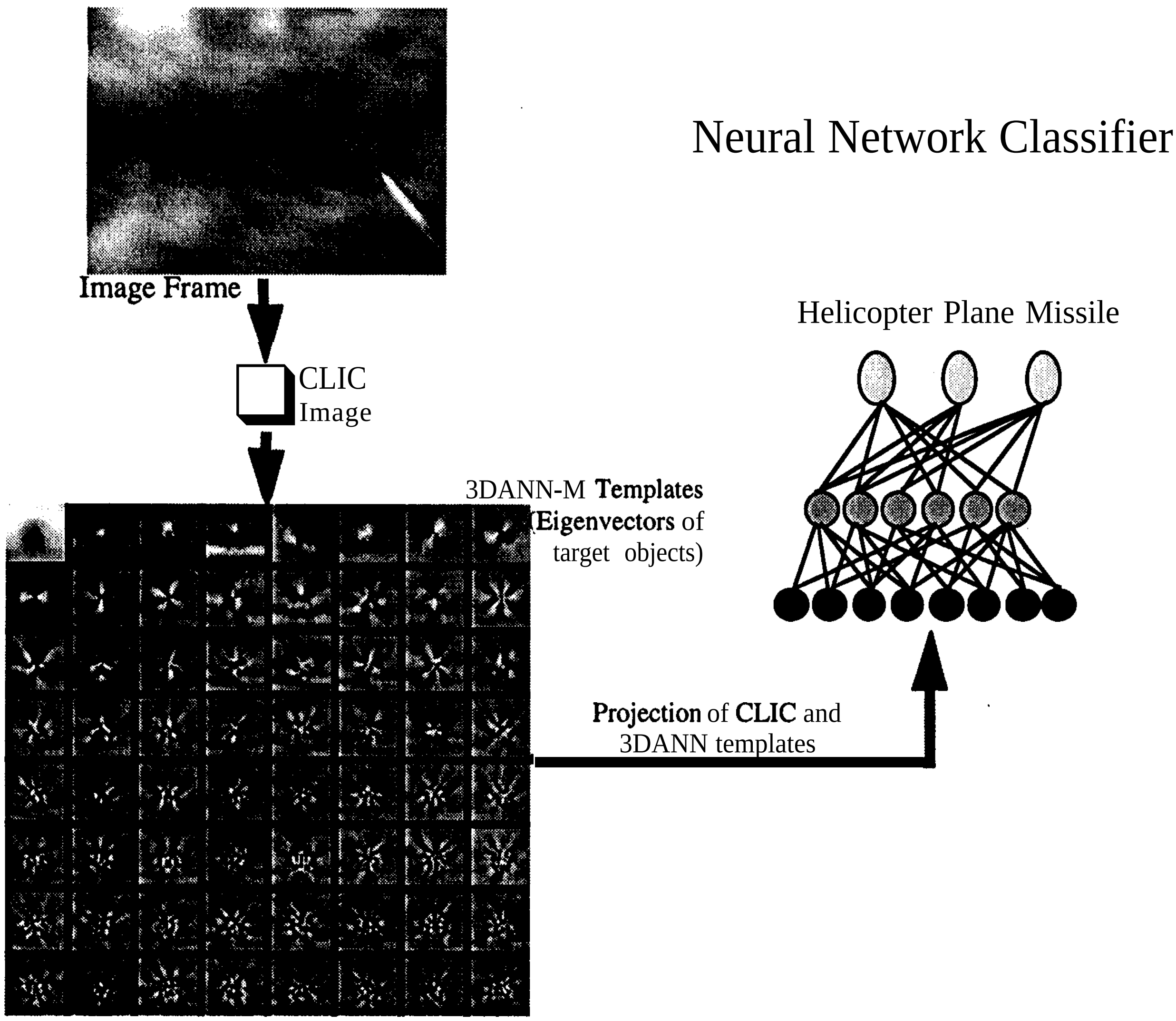


Figure 6. General target recognition is achieved using eigenvector projections in conjunction with a neural network classifier trained on selected data sets.

The most problematic aspect of this technique is that unless some restrictions are placed on variations in the target **imagery,** the most significant components become so general as to be unsuitable for fine distinctions such as object orientation or identity (e.g., missile type). Our strategy is to parametrize (e.g., lighting, pose, class, identity, and scale) and partition the object space in a hierarchical fashion, To classify each partition, a neural network (or other classifier) is trained on data imagery drawn from the set of variables that define the partition and projected onto eigenvectors suitable for that particular distribution of data.

Information about the object (its class, identity, or pose) is processed in a coarse-to-fine manner. For instance, after detecting an object in a **frame,** a rough estimate of image pose/scale is made, a result that can then **be** used to limit the variation that needs to be considered during object classification (i.e., plane, helicopter, and missile), Results using the technique described here have achieved nearly 97% detection rates, 94% classification rates for determining the angle of" the principle dimension of an object with respect to the image **(±30°),** and object classification rates approaching 95°/0. See Figure 7 for results on **object/nonobject** image classification rates achieved with a helicopter/missile/plane data set.

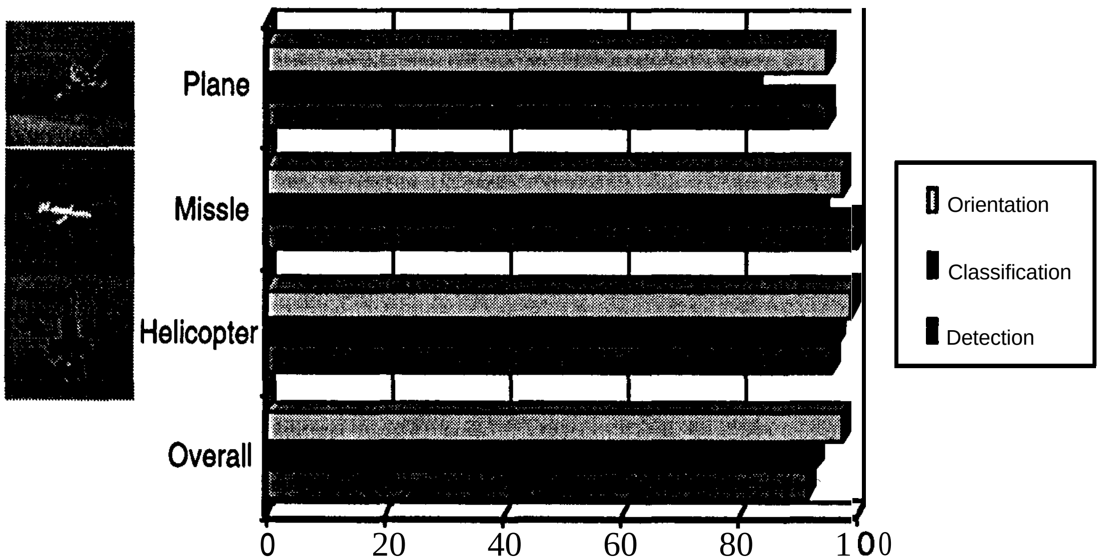


Figure 7. High detection/classification rates are achieved on selected data sets that include all possible orientations and scales of targets,

7. CONCLUSIONS

We have described a novel ATR system that uses advanced sensors and a combination of analog/digital processing modules in a compact, **efficient** package to achieve **difficult** object recognition tasks with unprecedented speed. Simulation studies have shown VIGILANTE's great flexibility for implementing complex hierarchical neural network algorithms. Further work is ongoing both in evaluating each sensor's effectiveness at the selected wavelength for BMDO targets-of-interest and in implementing an **ultrafast** end-to-end detection, classification, and precision-tracking methodology; both ground-baaed and airborne experiments using live targets are planned. The proposed hierarchical neural network algorithm also shows great promise for being able to achieve our real-time processing goal. By properly classifying the target and estimating scale/orientation in a hierarchical structure, precision tracking of selected target points (e.g., nose, wings, and tail) that uses standard model-based template-matching techniques can also be accomplished with great efficiency.

Future development would include integration of a two-color QWIP camera that covers both the medium-wave IR (MWIR) and long-wave IR (LWIR) spectra, enhancement of 3DANN-M to allow new functions (such as image "warping"), fabrication of the point-operation processor in silicon to replace CNAPS boards, and expanding the applications of VIGILANTE to commercial and space applications, Continuation of ATR methodology research and system verification using field experimental data are planned for 1997 through 1998.

8. ACKNOWLEDGMENTS

The authors would like to thank Dr. Sarath Gunapala of JPL for supplying Figure 2 for our use. The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the Ballistic Missile Defense Organization through an agreement with the National Aeronautics and Space Administration,

9. REFERENCES

1. B. Bhanu, "Automatic target recognition: state of the art survey," ***IEEE Trans. Aerosp. Electron. Sys.,*** vol. AES-22, no. 4, pp. 364-379, 1986.
2. D. Hammerstrom, E. Means, M. Griffin, G. Tahara, K. Knopp, R. Pinkham, and B. Riley, "An 11 million transistor digital neural network execution engine," ***Proceedings of the IEEE International Solid-State Circuits Conference,*** pp. 180-181, 1991.
3. D. K. Andes, J. C. Witham, and M. D. Miles, "MAVIS: A special purpose neural computational system for ATR," ***Neural Networks,*** vol. 8, no. 7/8, pp. 1349-1358, 1995
4, B. V. K. Vijaya Kumar, "Tutorial survey of composite filter designs for optical correlators," ***Appl. Opt.,*** vol. 31, no. 23, pp. 4773-4801, 1992.
5. B. V. Dasarathy, ***Decision Fusion,*** IEEE Computer Society Press, Los Alamitos, CA, 1993,
6. J. D. Mill and B. D. Guilmain, "The MSX mission objectives," ***Johns Hopkins APL Technical Digest,*** vol. 17, no, 1, pp. 4-10, 1996.
7. J. Carson, "On focal plane array feature extraction using a 3-D artificial neural network (3 DANN)," ***Proc. SPIE,*** vol. 1541, Part I: pp. 141-144, Part II: pp. 227-231, 1991.

8. T. Duong, S. Kemeny, T. Daud, A. Thakoor, C. Saunders, and J. Carson, "Analog 3-D neuroprocessor for fast frame focal plane image processing," *SIMULATION*, vol. 65, no. 1, pp. 11-24, 1995.
9. S. D. Gunapala, J. K. Liu, M. Sundaram, S. V. Bandera, C. A, Shott, T. Hoelter, P. D, Maker, and R. E. Muller, "Long-wavelength 256x256 QWIP hand-held camera," *Proc. SPIE*, vol. 2746, pp. 124-133, 1996.
10. E. R. Fossum, "Active pixel sensors: Are CCD's dinosaurs?", *Proc. SPIE*, vol. 1900, pp. 2-14, 1993.
11. S. Nikzad, M. E. Hoenk, P. J. Grunthaner, R. W. Terhune, R. Winzenread, M. Fattahi, H-F. Tseng, and M. Lesser, "Delta-doped CCDS as stable, high sensitivity, high resolution UV imaging arrays," *Proc. SPIE*, vol. 2278, pp. 138-146, 1994.
12. D. Levin, L. H. Caveny, and D. M, Mann, "Ultraviolet emissions quantified by rocket payloads," *Proc. SPIE*, vol. 1764, pp. 388-399, 1992.
13. W, S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115-133, 1943
14. T. Duong, S. Kemeny, M. Tran, T. Daud, and A. Thakoor, "Low power analog neurosynapse chips for a 3-D sugarcube neuroprocessor," *Proceedings of the IEEE International Conference on Neural Networks*, vol. III, Orlando, pp. 1907-1911, 1994,
15. T. Duong, T. Thomas, T. Daud, A. Thakoor, and B. Lee, "64x64 Analog input array for 3-dimensional neural network processor," *Proceedings of the 3rd International Conference on Neural Networks and Their Applications*, Marseilles, France, 1997.
16. M. Turk and A. Pentland, "Eigenfaces for recognition," *J. of Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.
17. C. Padgett, G. Cottrell, and R. Adolphs, "Categorical perception in facial emotion classification, " *Proceedings of the 18th Annual Conference of the Cognitive Science Society*, Hilldale, pp. 201-207, 1996.
18. C. Padgett, M. Zhu, and S. Suddarth, "Detection and object identification using VIGILANTE processing system", Proc. SPIE, vol. 3077, 1997.