NEUROMORPHIC PROCESSOR FOR REAL-TIME BIOSONAR OBJECT DETECTION

Gert Cauwenberghs[1], *R. Timothy Edwards*[2], *Yunbin Deng*[1], *Roman Genov*[1], *and David Lemonds*[3]

[1]ECE Department, Johns Hopkins University, Baltimore MD 21218
[2]Applied Physics Laboratory, Johns Hopkins University, Laurel MD 20723
{gert,tim,yunbin,roman}@bach.ece.jhu.edu

[3]Orincon Corporation, Kailua, HI 96734
lemonds@orinconhi.com

ABSTRACT

Real-time classification of objects from active sonar echo-location requires a tremendous amount of computation, yet bats and dolphins perform this task effortlessly. To bridge the gap between human-engineered and biosonar system performance, we developed special-purpose hardware tailored to the parallel distributed nature of the computation performed in biology. The implemented architecture contains a cochlear filterbank front-end performing time-frequency feature extraction, and a kernel-based neural classifier for object detection. Based on analog programmable components, the front-end can be configured as a parallel or cascaded bandpass filterbank of up to 34 channels spanning the 10 to 150 kHz range. The classifier is implemented with the *Kerneltron*, a massively parallel mixed-signal Support Vector "Machine" in silicon delivering a throughput in excess of a trillion (10^{12}) multiply-accumulates per second for every Watt of power dissipation. The system has been evaluated on detection of mine-like objects using linear frequency modulation active sonar data (LFM2, CSS Panamy City), achieving an out-of-sample performance of 93% correct single-ping detection at 5% false positives, and a real-time throughput of 250 pings per second.

1. INTRODUCTION

The bottle-nose dolphin excels at identifying buried objects from broadband sonar echo. The difficulty in attaining the recognition performance of biosonar in artificial systems arises from several factors, including strong variance of sonar returns under different orientations of the object, and partial occlusion and spurious echo by the sediment [1]. Recently, biosonar modeling combined with machine learning approaches have advanced significantly to realize automated real-time sonar classification systems [2, 3]. The key to practical deployment is efficient implementation of the algorithms in a compact low-power unit, suitable for integration in an underwater autonomous vehicle.

2. SYSTEM ARCHITECTURE

The approach we have taken is to map biosonar models of dolphin echo-location [1, 2, 3] onto a parallel architecture for dedicated implementation in hardware. Neuromorphic design principles adhere to biosonar models in functionality and structure. The architecture shown in Figure 1 consists of two main components: a cochlear front-end, and a neural classifier. The components are described next.

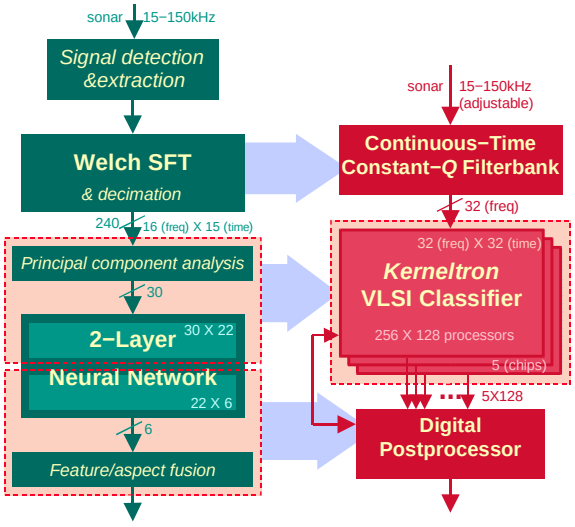

Fig. 1. *Architecture for sonar object classification.* Left: *Model of dolphin biosonar echo-location [1, 2].* Right: *Real-time hardware implementation, with reconfigurable front-end signal processor and massively parallel VLSI kernel-based classifier.*

2.1. Cochlear Front-End

A general-purpose continuous-time front-end filterbank is implemented for real-time extraction of time-frequency features in sonar signals. A cochlear model is adopted as a neuromorphic alternative to the Welch-windowed short-time Fourier transform typically used in biosonar models for target detection [1, 2]. A wide variety of cochlear models are supported, *e.g.*, using parallel [4] or parallel-cascade [5] filterbank topologies, for a maximum of 34 programmable frequency channels spanning the 10 to 150 kHz range typical of bat and dolphin sonar. Post-processing of the filtered signal produces time-frequency features as time-averaged

This work was supported by the ONR Biomimetic Broadband Biosonar Program, N00014-01-1-0315, N00014-99-1-0654, and N00014-99-1-0612.

amplitude-rectified spectra or pulse trains of zero-crossing events.

2.2. Principal Component Neural Classifier

The time-frequency features extracted from the cochlear front-end are presented as a vector to the neural classifier. To avert overfitting in training the classifier with 1,024 inputs from 32 frequency channels and 32 time decimation bins, principal components are extracted. Principal component analysis (PCA) is performed only during (off-line) training; the PCA layer and the input layer of the trained neural classifier are collapsed into one input layer of the run-time classifier implemented in hardware, since these two linear layers can be combined by multiplying the matrices.

This collapsed input layer is implemented by the *Kerneltron* [6, 7], a massively parallel VLSI neural network classifier loosely modeled after support vector machines [8]. The *Kerneltron* implements a fully connected synaptic layer with 512 input units (or vector components) and 128 output units (or support vectors), for a total of 65,536 programmable connections. An array of four *Kerneltron* chips are embedded on a printed circuit card interfacing with a host PC and the front-end processor, for a total of 1,024 input units and 128 hidden neurons (support vectors). Since inputs represent time-frequency features, each output implements a cortical cell with particular time-frequency response as observed in the mammal auditory system [9] and implemented in silicon models of acoustic transient classification [10, 11]. These cortical units are combined in the output layer for object recognition.

3. IMPLEMENTATION

Flexibility is maintained in tailoring algorithms to the biosonar application by allowing for programmable system parameters and reconfigurable topology in the implemented architecture.

3.1. Front-End Biosonar Processor

The filterbank front-end is implemented using analog circuits based on the Lattice Semiconductor ispPAC family of field-programmable analog chips. Reconfigurability in the architecture is provided through non-volatile, re-writable configuration memory in the filter and interface chips. Filter topology (bandpass, lowpass) and parameters (center frequency, quality factors) are programmable through the PC interface. Inputs may be selected from a number of different sources, including direct analog sonar input. Outputs are digital and are presented in serial fashion for direct interfacing with the *Kerneltron*.

3.2. The *Kerneltron*: Massively Parallel VLSI Synaptic Array

The *Kerneltron* processor [6, 7] is internally analog and externally digital, combining the best of both worlds: the efficiency of analog computing, and the flexibility of a reconfigurable digital interface. The core contains an array of CID/DRAM cells, unit memory cells each with a dedicated processor performing binary multiplication and analog accumulation. The massively parallel architecture with fine-grain distributed memory and computation enables an energetic efficiency (operations per unit energy or, equivalently, throughput per unit power) and integration density (throughput per unit area) a factor 100-10,000 better than that of high-end multiscalar processors or DSP chips [6]. A prototype of the *Kerneltron* VLSI chip, comprising the core of the classifier architecture, is shown in Figure 2.

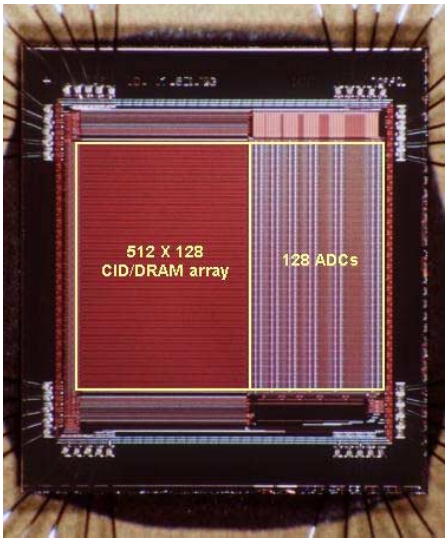

Fig. 2. *The* Kerneltron, *a massively parallel VLSI array processor for kernel-based pattern recognition (Genov and Cauwenberghs, 2001 [6]).*

With the present implementation, 8-bit of digital resolution is obtained for each support vector output. Higher precision, unlimited by the analog system precision, can be attained using stochastic encoding of the digital input [7]. Newer generation *Kerneltron* chips have been designed and are presently being incorporated in the system.

3.3. Configurability and Scalability

Other parts of the architecture, including the output layer of the neural network combining the *Kerneltron* outputs, are implemented using a Xilinx FPGA (field programmable gate array) with supporting software, for maximum flexibility and programmability. This solution is viable since the outputs from the Kerneltron are relatively low data rate, and optimized parallel hardware is not desirable for this part of the architecture. The FPGA also interfaces with the Kerneltron to configure system parameters in the implemented kernel, and perform digital post-processing and resolution enhancement for optimal performance.

Extensions on the architecture can be accommodated in fairly straightforward fashion, since the design parameters (number of time-frequency bins, number of cortical ATP cells, etc.) can be expanded in modular and incremental fashion by tiling together processors in parallel.

4. REAL-TIME MINE DETECTION: EXPERIMENTAL VALIDATION

The programmable nature of the architecture implies that the system can be used for a variety of tasks. To assess biosonar classification performance, we used benchmark LFM2 data from the Navy Coastal Systems Station, Panama City, FL. The LFM2 data contains backscatter from 2 (bullet, manta) mine and 4 (drum, limestone, granite, log) non-mine objects, for linear frequency modulation broadband sonar covering the 15-150kHz range. Each object is ensonified at 72 different angles with 5 degree intervals.

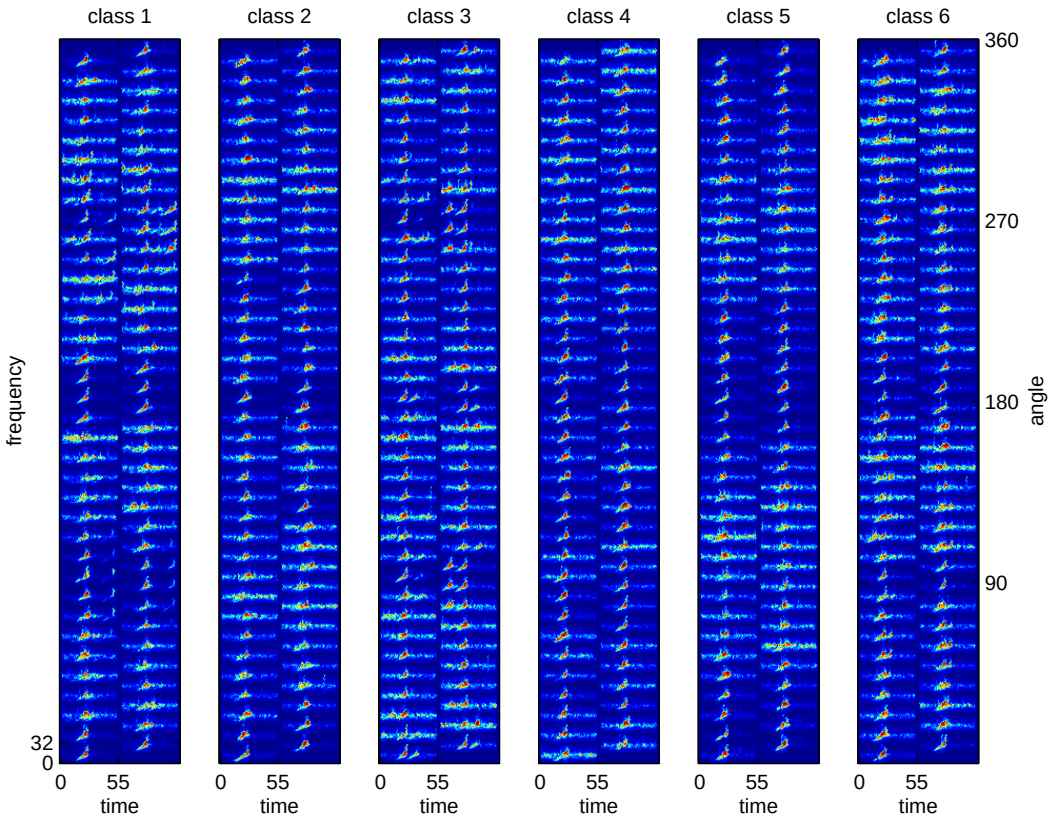

Fig. 3. *Experimental time-frequency feature outputs obtained in real-time from the biosonar signal processing front-end, on the LFM2 data. Training and test data are distributed over alternating angles (left and right columns) for each target.*

For training and test, the data is separated in equal parts according to alternating angles for each object. A parallel constant-Q bandpass filterbank topology [12] is used for the cochlear frontend with 32 center frequencies linearly spaced from 15 to 150kHz. Object labels define 6 classes for support vector machine supervised learning.

Proof-of-concept results that demonstrate training and classification of sonar data on the system in real-time are shown in Figures 3 and 4. The ROC (receiving operating characteristic) curve for a mine/non-mine classification task on the LFM2 data is shown in Figure 5. The hardware returns an out-of-sample recognition performance of 93% correct detection at 5% false positives, while the simulations in software return 95% correct detection at the same false positive rate. The hardware results run in real-time (2 msec per ping), presenting a 200-fold improvement in computation speed over the software running on a Pentium III workstation.

The 2% performance gap between hardware and simulated detection rates is attributed to the frequency range of the biosonar front-end which borders the range of the returned sonar signal, and the finite precision of programmed system parameters of the front-end resulting into differences in software models and implemented hardware. Both can be redressed in a dedicated VLSI version of the front-end [12].

5. CONCLUSIONS

This effort fills a gap in computational power for real-time biosonar classification and, ultimately, the practical realization of autonomous mine detection systems embedded in small underwater vehicles. The hardware offers a factor 200 improvement in computing performance (pattern classifications per second) over general-purpose compute, running in real time. The potential of integrating the system in a low-weight miniature and micropower package makes it especially attractive for use in an unmanned underwater vehicle.

Substantial gains in recognition performance can be expected from embedding the hardware in a system for acquisition and preprocessing of massive amounts of biosonar data in a real-world environment. Large amounts of data are needed for training a sonar classifier to be relatively insensitive to variations in the environment such as sediment, background clutter, and variations in the actual target. We envisage that the tool presented here will enable automated on-line acquisition and preprocessing of a continuous stream of data.

6. REFERENCES

[1] Roitblat, H.L., W.W.L. Au, P.E. Nachtigall, R.H. Shizumura, and G.C. Moons, "Sonar Recognition of Targets Embedded in Sediment," *Neural Networks*, vol. **8**, pp. 1263-1273, 1995.

[2] Shizumura, R.H., W.W.L. Au, G.C. Moons, P.E. Nachtigall,

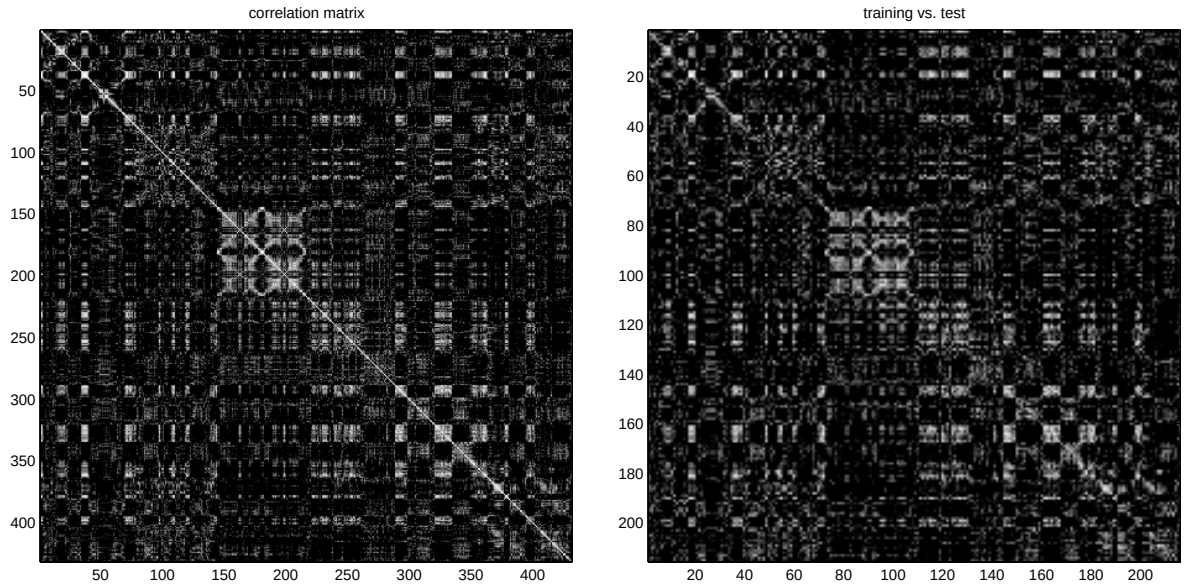

Fig. 4. *Support vector training of experimental LFM2 data. (a): Correlation matrix, and (b): training vs. test cross-correlation matrix. The support vector machine is constructed using kernels derived from the partial training vs. training correlation matrix.*

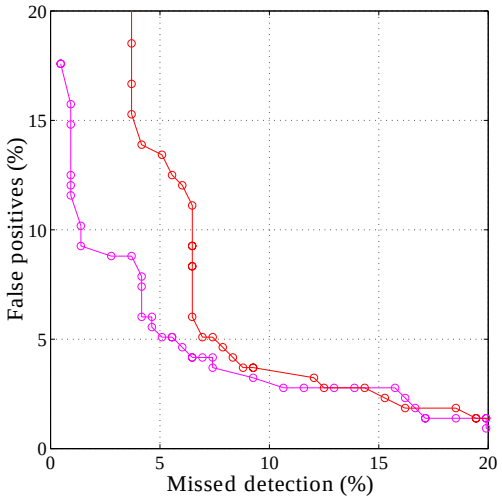

Fig. 5. *Real-time vs. emulated sonar classification. ROC (receiver operating curve) for mine-like vs. non-mine classification on the test set, after training on the simulated and measured training data.*

H.L. Roitblat and R.C. Hicks, "Real-Time Sonar Classification of Proud and Buried Targets Using Dolphinlike Echolocation Signals," *J. Acoustical Soc. Am.*, vol. **102**, p. 3152-, 1997.

[3] G.S. Okimoto, R. Shizumura, and D. Lemonds, "Active biosonar systems based on multiscale signal representations and hierarchical neural networks," in *Detection and Remediation Technologies for Mines and Minelike Targets III, SPIE Proc.*, vol. **3392**, pp.316-323, 1998.

[4] L. Watts, D.A. Kerns, and R.F. Lyon, "Improved Implementation of the Silicon Cochlea," *IEEE Journal on Solid-State Circuits*, vol. **27** (5), pp 692-700, 1992.

[5] W. Liu, A.G. Andreou, and M.G. Goldstein, "Voiced-Speech Representation by an Analog Silicon Model of the Auditory Periphery," *IEEE T. Neural Networks*, vol. **3** (3), pp 477-487, 1992.

[6] Genov, R. and G. Cauwenberghs, "Charge-Mode Parallel Architecture for Matrix-Vector Multiplication," *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. **48** (10), Oct. 2001.

[7] Genov, R. and G. Cauwenberghs, "Stochastic Mixed-Signal VLSI Architecture for High-Dimensional Kernel Machines," *Adv. Neural Information Processing Systems (NIPS'2001)*, Cambridge: MIT Press, vol. **14**, 2002.

[8] Vapnik, V., *The Nature of Statistical Learning Theory*, Berlin: Springer-Verlag, 1995.

[9] Wang, K.S., and S.A. Shamma, "Spectral Shape Analysis in the Central Auditory System," *IEEE T. Speech and Audio Processing*, vol. **3** (5), pp. 382-395, 1995.

[10] Pineda, F.J., G. Cauwenberghs, R.T. Edwards, K.T. Ryals and D.G. Steierwald, "Bang, Click, Thud, or Whack?," *Johns Hopkins APL Tech. Digest*, vol. **18** (2), pp. 244-253, 1997.

[11] Edwards, R.T., and G. Cauwenberghs, "Mixed-Mode Correlator for Micropower Acoustic Transient Classification," *IEEE J. Solid-State Circuits*, vol. **34** (10), pp. 1367-1372, 1999.

[12] Edwards, R.T., and G. Cauwenberghs, "Synthesis of Log-Domain Filters from First-Order Building Blocks," *Analog Integrated Circuits and Signal Processing, Int. J.*, vol. **22**, pp. 177-186, 2000.