The First Bi-directional Neural Network

A Device for Machine Learning and Association for a smarter, faster, more agile, and more transparent Human-Computer Interface

Technical Presentation and Discussion with

SSC PAC

by

James P. LaRue

September 18 2013

Outline

- Slide 3 – History
- Slide 4 – Take the Edge off Pure Logic
- Slide 5 – The Bi-directional Neural Network and HCI
- Slide 6 – Results
- Slide 7 – Discussion Topics
- Slide 8 – Thank you.   Contact Information

- Backup Slides 9-12 – Credits, Why it's fast, More Results, Chalkboard Ideas
- Plus slides 13-15 for AVIPE

Historical Path to the Bi-directional Neural Network

1. Convolutional Neural Network - CNN

Alexander Bain (1873) and William James (1890) *Neurons interact.*

McCulloch and Pitts (1943) *1st computational model.*

Rosenblatt (1958) *Feed-forward perceptron, convergence issue.*

Werbos (1975) *Fixes the Rosenblatt problem, goes unrecognized.*

Fukushima (1980) *Neocognitron – hidden layer visual pattern recognition.*

Rumelhart, Hinton, Williams, McClelland (1984) *Recognize Werbos work.*

LeCun and Bengio (1995) Convolutional '*Neocognitron' (CNN), long training*.

2. Associative Memory Matrix – AMM

Kosko (1988) *Bi-directional I/O matrix, no hidden layers, stability issue.*

3. Couple of Comments

Minsky and Papert (1969) *Need at least one hidden layer between I/O to be meaningful.*

Cybenko/Hornik (1989) *Universal Approximation Theorem (UAT), one single hidden layer.*

4. Bidirectional Neural Network - BNN

Luzanov/AFRL /AFOSR(2011) Create a bi-directional model of ventral (vision) pathways.

LaRue (2012) Translated CNN inter-layers into bi-directional AMM structure.

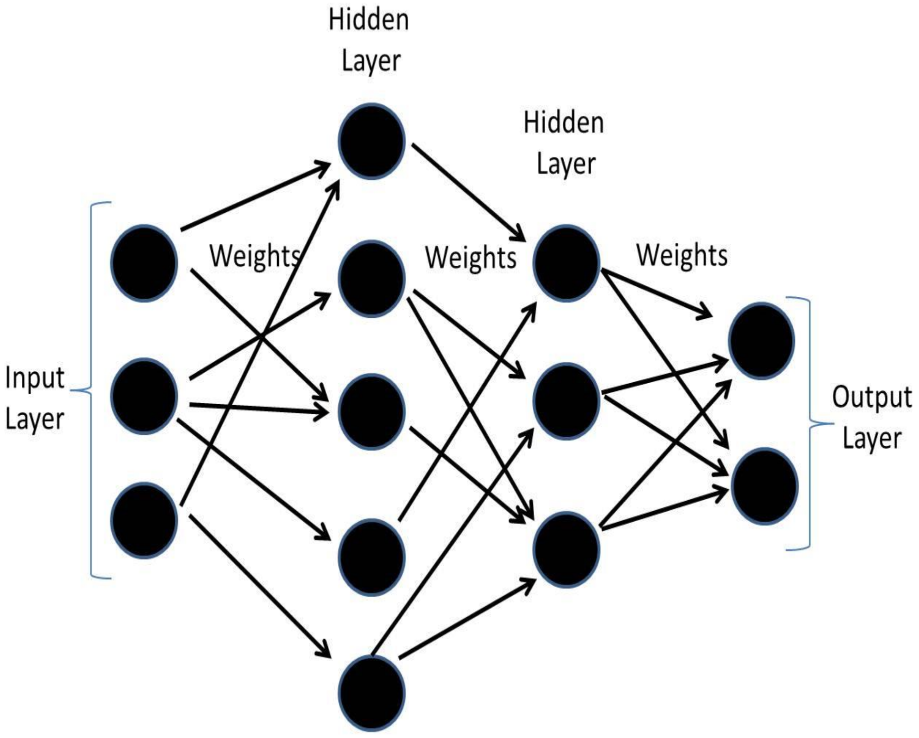
LaRue (2013) Met UAT criterion, formed intra-layer connections, mutual benefit.

Result: Smarter training, faster execution, Inter/Intra communication

Bring to Machine Logic an Element of Machine Intuition

Neural Network

weights and neurons
connecting I/O



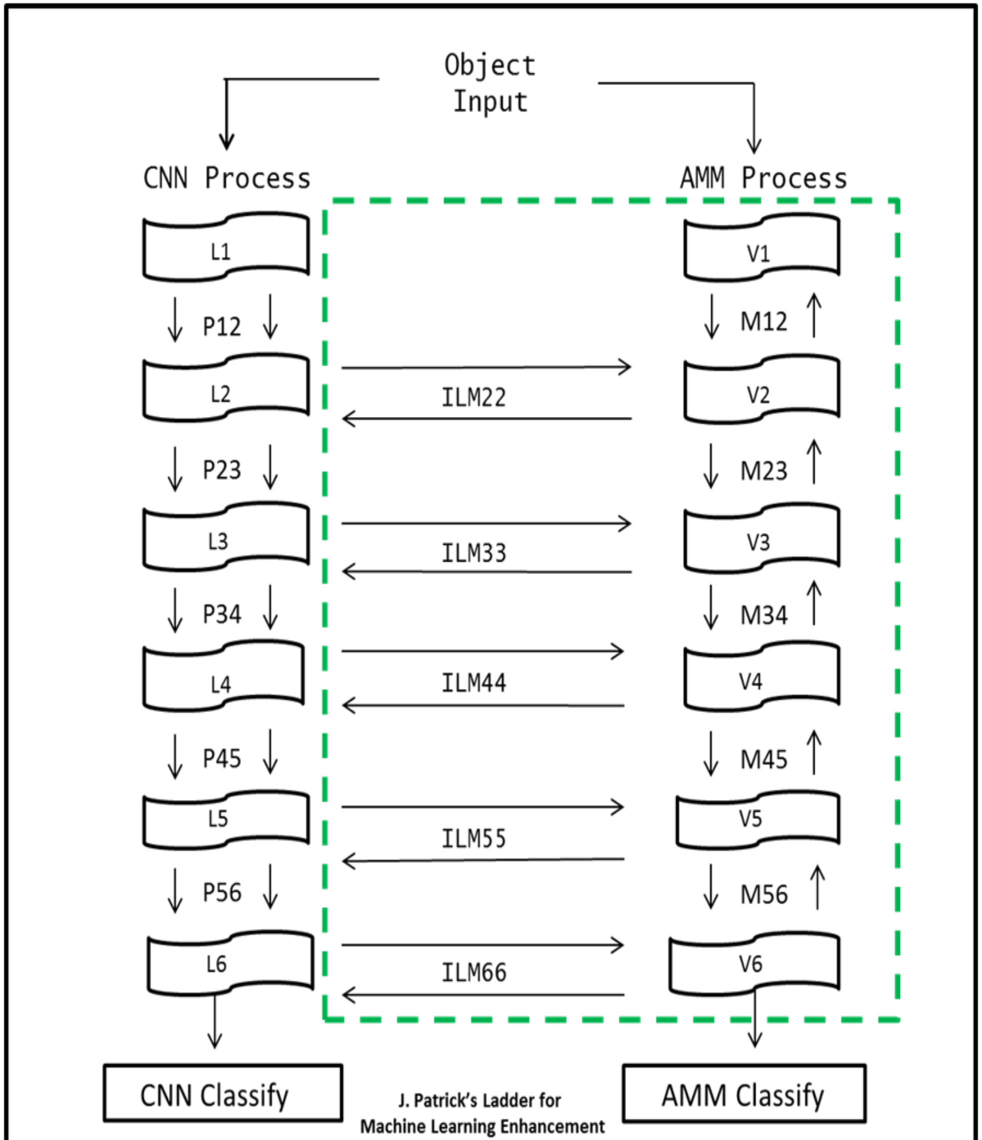
Associative Memory Matrix

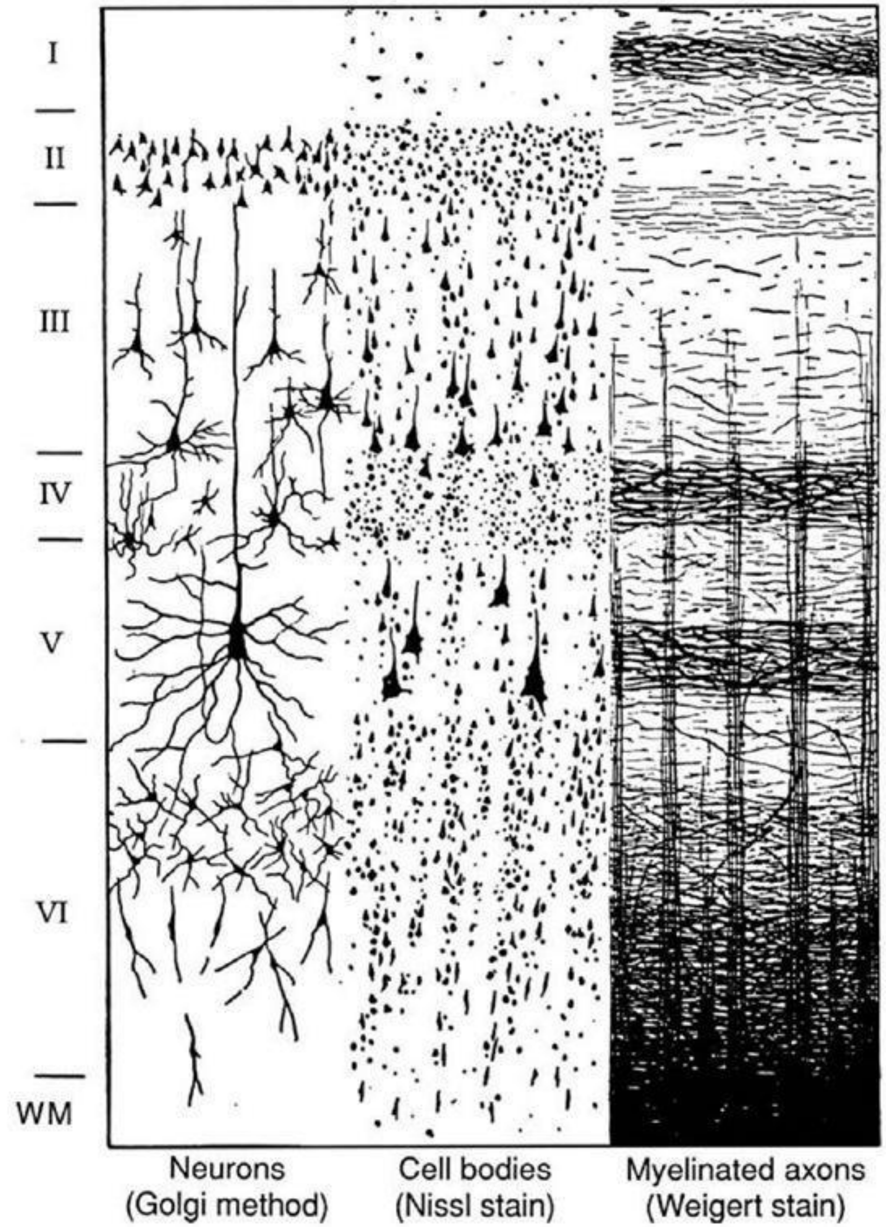
I/O outer products
connecting I/O

INPUT	OUTPUT
X1 = (1, -1, 1, -1, 1, -1)	Y1 = (1, 1, -1, -1)
X2 = (1, 1, 1, -1, -1, -1)	Y2 = (1, -1, 1, -1)

$$M = \sum X_i^T \, Y_i = \begin{bmatrix} 2 & 0 & 0 & -2 \\ 0 & -2 & 2 & 0 \\ 2 & 0 & 0 & -2 \\ -2 & 0 & 0 & 2 \\ 0 & 2 & -2 & 0 \\ -2 & 0 & 0 & 2 \end{bmatrix}$$

The Bi-directional Neural Network and HCI

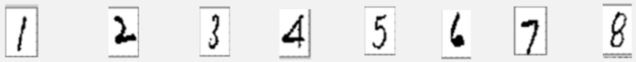




J. Patrick's Ladder
Customer Data Test Sets

Customer	Data Type	Learning Process
Air Force Research Laboratory & Air Force Office Scientific Research	Hand Written Numerals MNIST Data Set	CNN + Perceptron
Neurotechnology	Biometric Data Fingerprints	Perceptron Only
Defense Advanced Research Projects Agency	Image Data Armed Personnel	CNN + Perceptron
Pennsylvania State & Applied Research Laboratory	Video Data Person with Object/Weapon	CNN + Perceptron

MNIST Numerals



Person with Object/Weapon

<table>
<tr><th colspan="4">J. Patrick's Ladder Results</th></tr>
<tr><td rowspan="3">CNN + Perceptron Cases</td><td>Hand Written Numerals</td><td>Training</td><td>Execution</td></tr>
<tr><td>Armed Personnel</td><td rowspan="2">10X faster</td><td rowspan="2">20X faster</td></tr>
<tr><td>Person with Object/Weapon</td></tr>
<tr><td rowspan="2">Perceptron Only Case</td><td rowspan="2">Fingerprints</td><td>Training</td><td>Execution</td></tr>
<tr><td>10X faster</td><td>3X faster</td></tr>
</table>

Discussion

- **Can ML recognize relevant objects from imagery?**

✓ Take an SSC-PAC ML algorithm from computer vision convert to BNN architecture, validate, 200 hours.

✓ Submit joint patent for J. Patrick's Ladder. This BNN is a innovation, first of its kind, a 25 year break-through. 200 hours.

✓ Form team for ONR FY2014 MURI TOPIC #19, Role of Bidirectional Computation in Visual Scene Analysis – PMs: Harold Hawkins and Behzad Kamgar-Parsi wrote: *…almost all visual cortex models are based on feed-forward projections, …although, it is well known that neural connections in biological vision are bidirectional.*

- **Can ML recognize relevant MSG traffic based on changing context?**

✓ (1) Strategic: MSG traffic as neuron pulse. (SONAR for the Internet).

✓ (2) Tactical: NLP with AMM.

- **Can autonomous vehicles learn new tasks with limited user instruction?**

✓ Reverse of (2) NLP with AMM.

✓ Is any one using both eyes? (Get 40% cross-over).

- ***How can humans and AI/ML work together to create better analysis results?***

✓ For starters, a bi-directional communication framework. Top-Down/Bottom-Up.

Thank you SSC PAC

For more information on the
Bi-directional Neural Network
for Biological and Man-made Systems

Contact: James LaRue, PhD
James@jadcoSignals.com
www.JPatricksLadder.com
www.JadcoSignals.com
315 717 9009

CREDITS

Adam Bojanczyk – Cornell – Extended Matrix Methods

Catalin Buhusi – Medical University of South Carolina – Striatal Beat Frequency in Axon firings

Ron Chapman – Nunez Community College – History of The Additive Model

Graciela Chichilnisky – Columbia University – Black Swan Theory

Leon Chua - Berkeley - Memristor (Nano)Technology

Bill Copeland – DARPA Innovation House – Clutter analysis

Yuriy Luzanov – AFRL RIGG - Working CNN algorithm and PM for BAM

Angel Estrella – University of Yucatan – Local Stability –June 28 at Griffiss Institute

Stephen Grossberg – Boston University – The Additive Model

Lauren Huie – AFRL RIEC/Penn State Grad – Diversity and vestiges of SVD in Nullspace Identification

Randall King – Avondale Shipyards – RF Waveform Analysis

Aurel Lazar – Columbia University – Neuromorphic Time Encoding Machine

Scott Martinez – SUNYIT Grad – RANDU and the Chinese Remainder Theorem

Todd Moon – Utah State University - Mathematics of Signal Processing (Great Book)

Louis Narens – University of California – Non-Boolean Algebra and bounded sequences

Kenric Nelson – Complexity

Andrew Noga – AFRL Information Directorate – Signal Processing

Mark Pugh – AFRL Information Directorate – Image Processing

Tomaso Poggio - MIT CBCL - HMAX

Edmond Rusjan – SUNYIT – Fourier Transform, Matrix Methods, and Sequences

George Smith – NRL/University of New Orleans – Multipath/ G-ILETS

Richard Tutwiler – Penn State – ICA and Learning Algorithms

Alfredo Vega – AFRL RIEC – Linear Recursive Sequences

Andy Williams - AFRL-RIEC - DCPs: SERTA and SCORE

James LaRue – University of New Orleans and *JadcoSignals* – Combined the ideas to form BNN

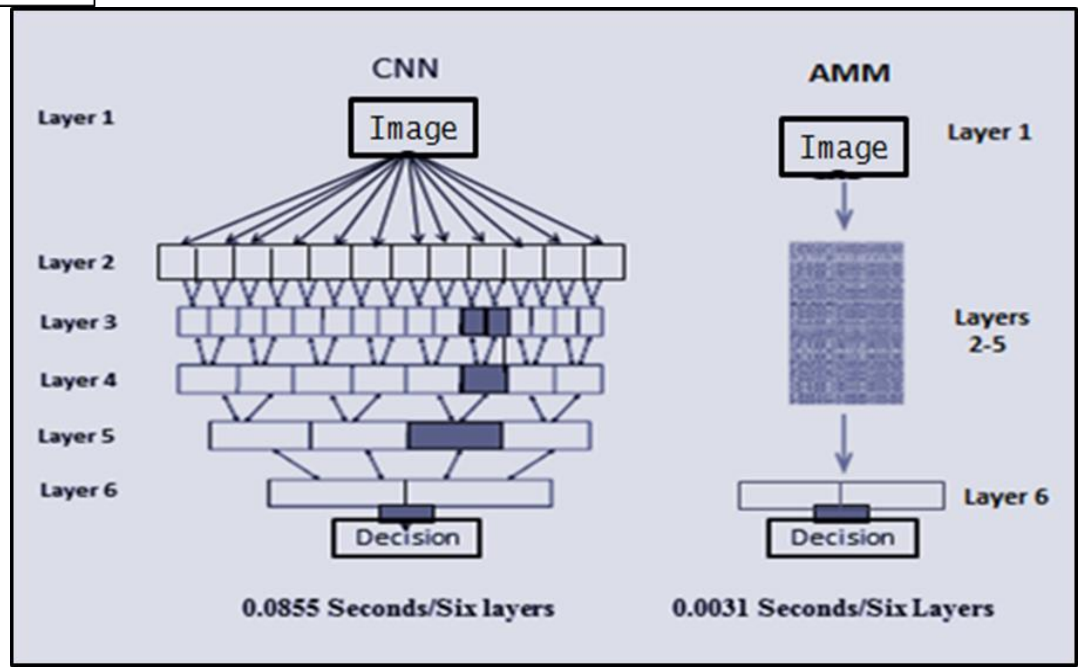
Connecting two layers

CNN		AMM	
Image IN L1	28x28	Image IN V1	784x1
1. 12 KV1 Filters		1. AMM	784x2352
2. 2D Convolution			
3. Tanh compress			
4. Down sample by 4			
5. Multiply by weights			
6. Add biases			
7. Tanh compress			
L2 out	(12) 14x14	V2 out	2352x1

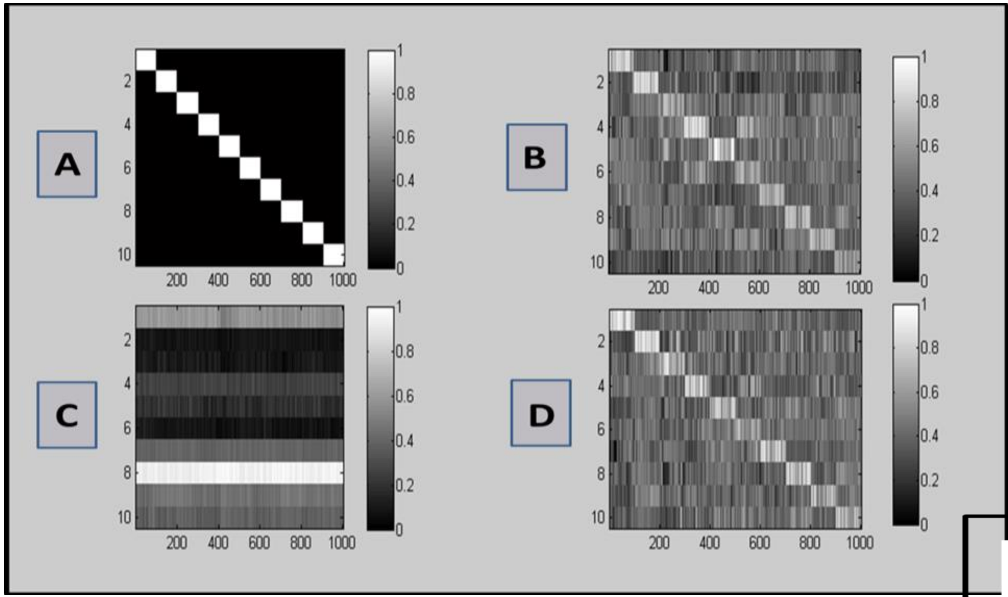
From seven steps to one step.

THE 1st AMM advantage

From four hidden layers to one hidden matrix.



Results



<u>The 1st AMM advantage is: *Speed*</u>

Also,

AMM is consistent , offers diversity

A IDEAL

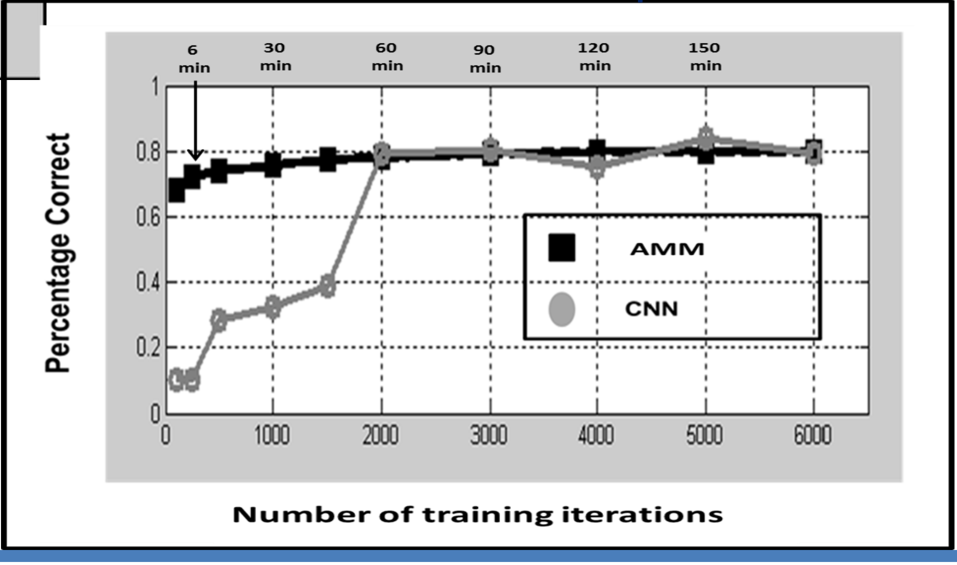
B. CNN **(0.0855 sec/decision)**

C. Unstable AMM

D. Stable AMM **(0.0031 sec/decision)**

Goal: Ideal staircase. D matches B accuracy.

6000 Iterations = One epoch



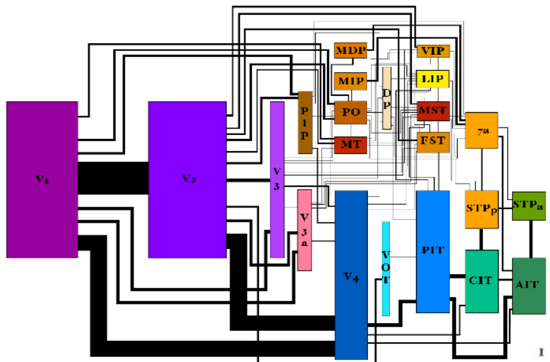
<u>The 2nd AMM advantage is: *Less Training*</u>

- AMM Claimed 97% of accuracy 10x faster
- BNN boosts CNN accuracy from 12% to 37% at 6 min mark

BIGGER PICTURE: From a cognitive science point of view, the BNN combines the logic-based neural network with the intuitive-based associative memory, resulting in a beneficial, bidirectional, inter-action and intra-action of diverse, yet complimentary, thought process.

Hubel &Wiesel, Rosenblatt, Fukushima, Poggio

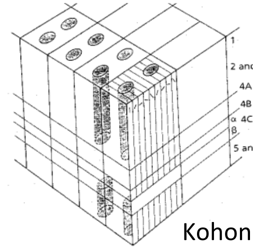
The ventral stream

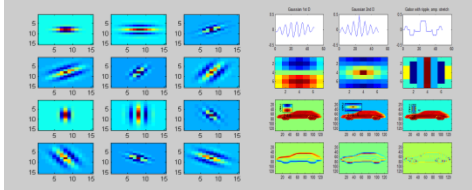


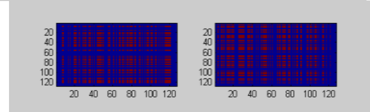
Source: Lennie, Maunsell, Movshon

Jeff Hawkins

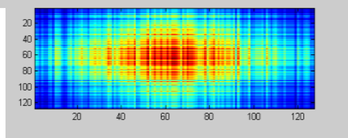
Hierarchical Temporal Model



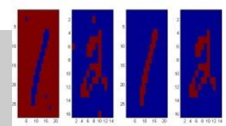




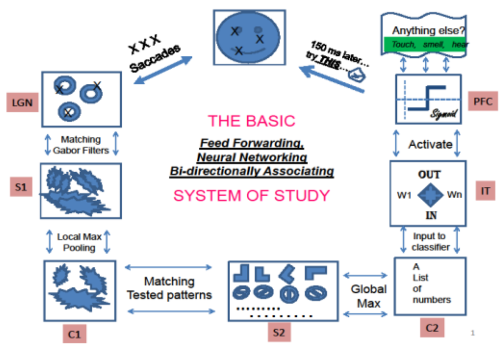
G I-lets

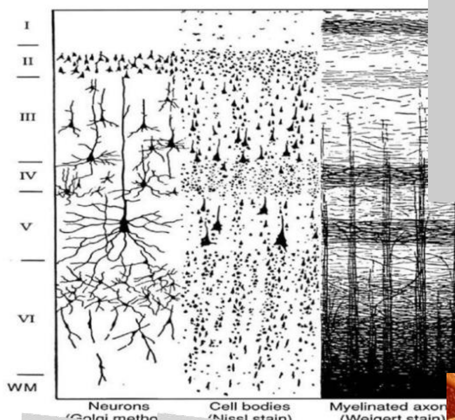


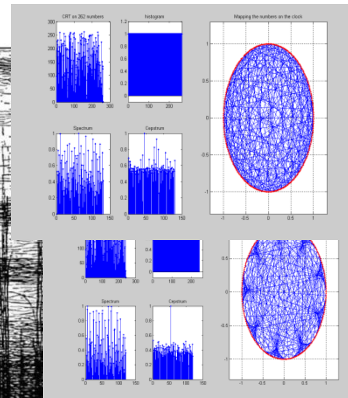
Kohonen, Kosko,



Cohen-Grossberg, Hopfield

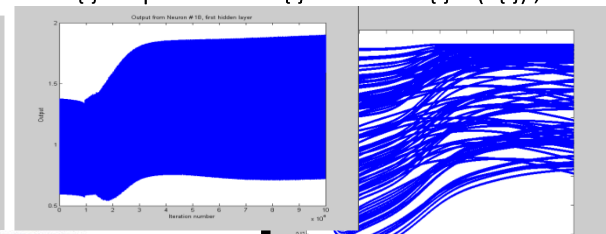


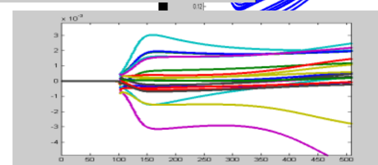


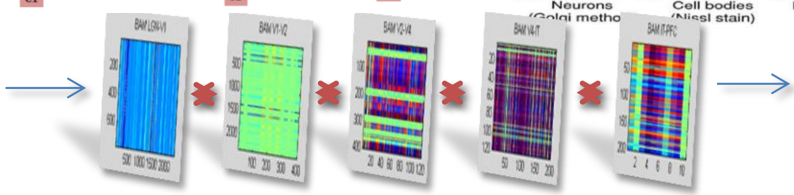


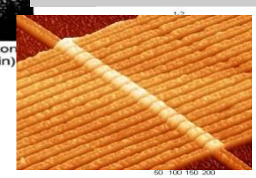
Widrow, and Adeline, Werbos and back-propagation

deltaw{l} = alpha * deltaw{l} - mu*delta{l} * (Y{l})';

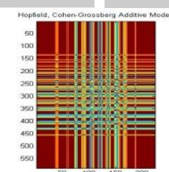






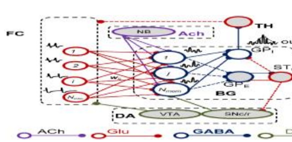


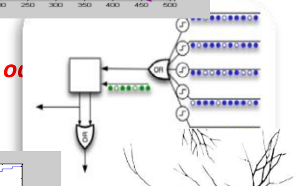
Leon Chua, Memristors

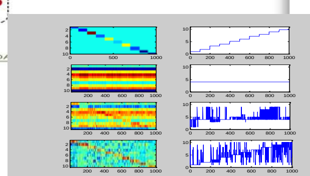


Striatal Beat-Frequency o

Meck, Buhusi,



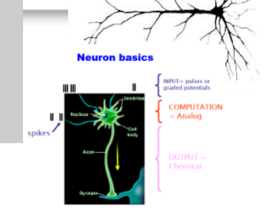




Louis Narens, Support Theory Based on a Non-Boolean Event Space, *need not satisfy the principles of the Law Of The Excluded Middle and the Law of Double Complementation.* **Graciela Chichilnisky** extend the foundation of statistics to integrate rare events that are potentially catastrophic, called ***Black Swans***.

Richard Tutwiler, Kenric Nelson, Edmond Rusjan, Scott Martinez, Adam Bojanczyk, Randall King, Mark Pugh, Andrew Noga, Ron Chapman, Bill Copeland, Angel Estrella – University of Yucatan, Alfredo Vega, Lauren Huie, Hugh Williamson, Andy Williams, Yuriy Luzanov, Jay Myung, Mike Geertsen

James P. LaRue *dba JadcoSignals* – Combined their ideas to form BAM, Philosophically speaking



Aurel A. Lazar

Neuromorphic model of spike processing

AFRL/AFOSR Cognition and Decision Program 2010-2012

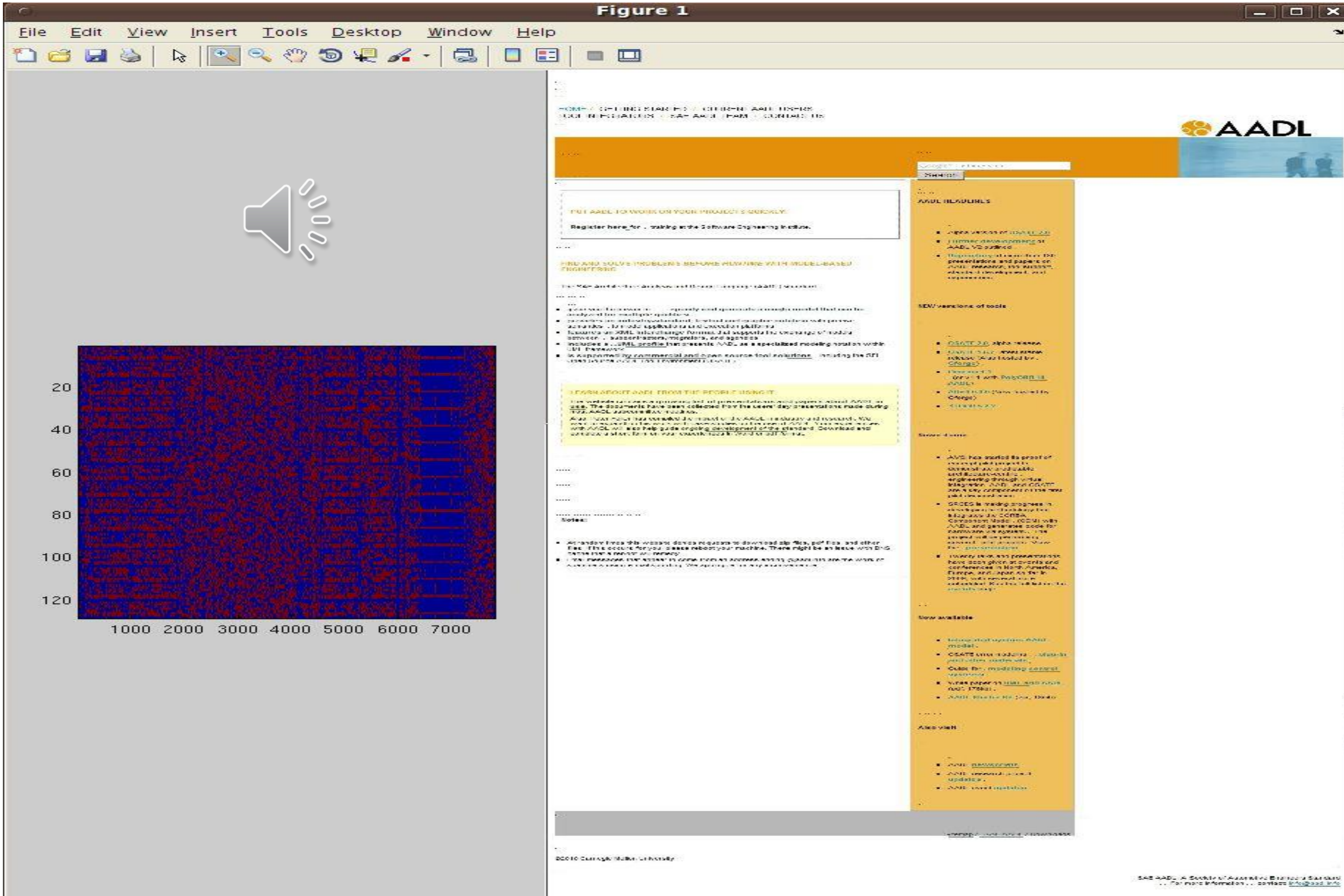
DARPA Innovation House Sept-Nov 2012

Jadco Signals 2011-

twentyoneseconds

www.DataPlasticity.com

All previous and:, “Cars.jpg” image (not displayed by html)
Men walking image



Treating Data Overload from a Speech Processing Point of View

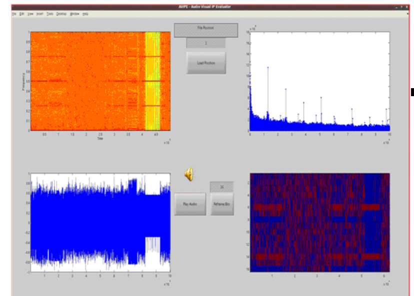
OBJECTIVE

Develop a real-time automated Intel data screening process to identify information areas of interest in intelligence network traffic using cepstral-based analysis techniques that will provide the intelligence analyst with audible and visual aids integrated with GV™ 3.0 operational viewer application as a plug-in module.

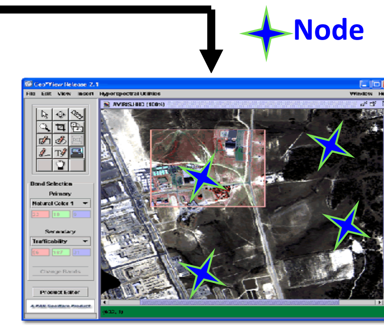
EXPECTED RESULTS/BENEFITS

The resultant Audio Visual IP Evaluator (AVIPE) will provide a method for screening real-time intelligence network traffic and data flows for key elements of information that can then be thoroughly analyzed for intelligence content. AVIPE will produce a new data processing paradigm that will increase efficiency by reducing data intensive operations in screened areas that don't warrant further analysis while directing attention to those areas of information that do.

AVIPE GUI



Integrated as Plug-in Module to GOTS Operational Viewer



This is a sample of the GV™3.0 display. In this screenshot, the color bands that display the trafficability of the selected terrain have been enhanced.

Measures the Intel Pulse of Network Traffic nodes(s) in Real-Time

GV™ 3.0

PERFORMANCE METRICS

	SOA	Advancement
Metric 1	Non-real-time key word matching schemas for archived data searches	Real-time analysis of dynamic network traffic intelligence data
Metric 2	No data reduction	Filtering yields 1000:1 reduction in streaming traffic
Metric 3	Single command driven analysis based on serial searches	Ability to partition Intel data flows into regions based on the cepstral coefficients and alert other platforms to search for matching or supporting Intel regions

PERFORMANCE OBSTACLES

The technical risks associated with this effort are deemed to be moderate in nature. The perceived risk is related to being able to tune the cepstral coefficients such that the process will be able to detect the information areas of interest and that the detection is reliable and repeatable. Acquiring and utilizing a reasonable but representative data set will ensure the opportunity to assess the overall AVIPE capability.

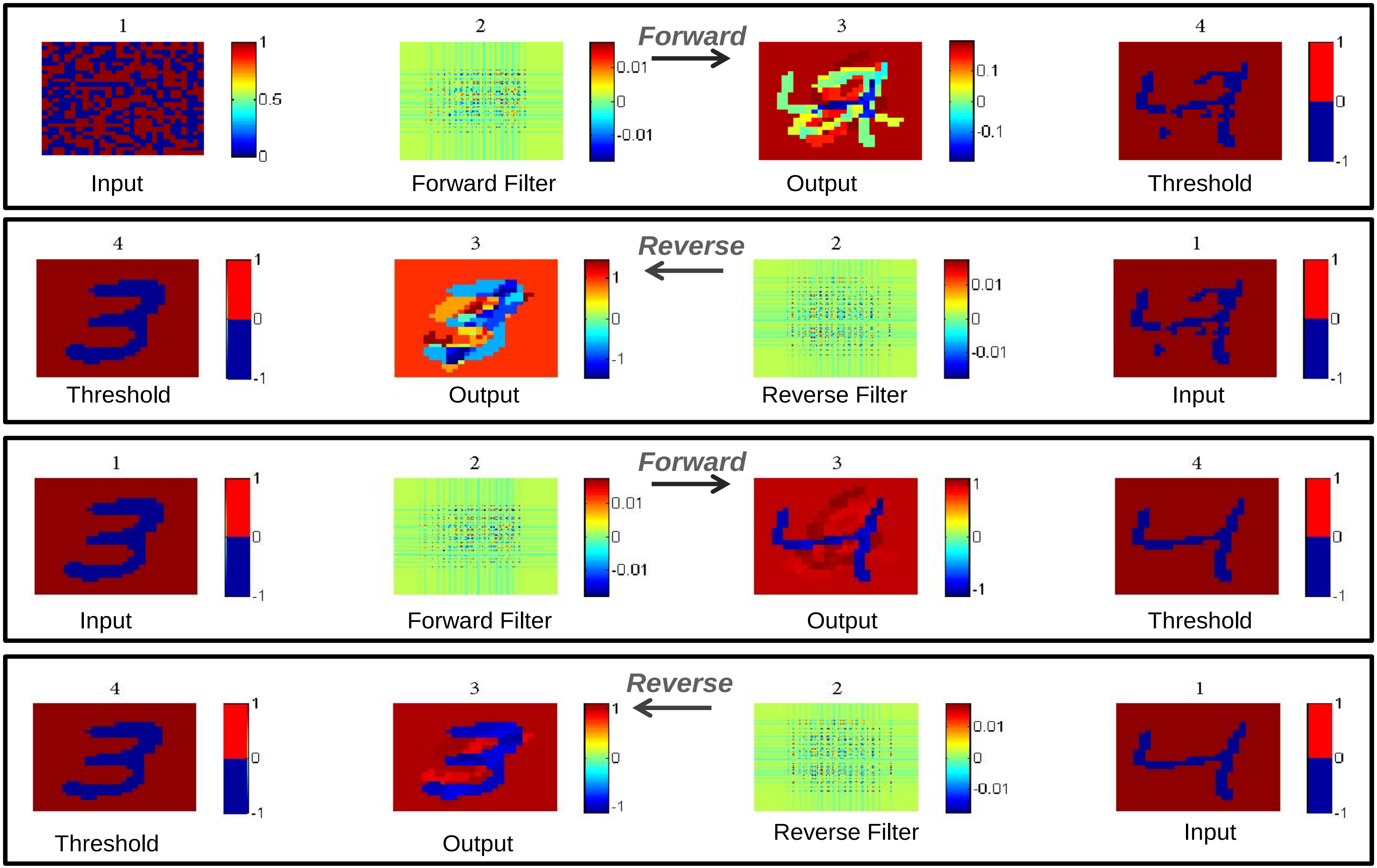
TECHNOLOGY READINESS LEVEL

Effort Start: TRL-2 Effort Completion: TRL-6

AVHCI - CCoF

Bidirectional Resonance

Formed Pairs: (1-2) (3-4) (5-6) (7-8)



Processing Time Breakdown

Function Name	Calls	**Total Time**	Self Time*	Total Time Plot (dark band = self time)
ZmyCNNTestB	1	118.042		

Line Number	Code	Calls	Total Time
84	convresultsV3(i).data = conv2(...	497000	20.296 s
23	convresultsV1(i).data = convre...	12000	13.339 s
83	de=KV3(connlistV3(i).list(j),i...	497000	12.510 s
82	sw=IMs(connlistV3(i).list(j))....	497000	12.234 s
51	convresultsV2(i).data = convre...	60000	8.380 s
All other lines			51.283 s
Totals			118.042 s

Function Name	Calls	**Total Time**	Self Time*	Total Time Plot (dark band = self time)
BAMTest BAMF	1	1.200 s	1.200 s	