Proposal Abstract

In Response To: **BAA 93–06**
Technical Topic Area: **Monolithic Sensors**

ALVINN-On-A-Chip: A Trainable Single-Chip Solution for Image-Motor Control

Technical Contact

Takeo Kanade
EMail: tk@cs.cmu.edu
FAX: 412-621-1970
School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3891

Administrative Contact

Maryann Brendel
EMail: mb4l+@cs.cmu.edu
Office of Sponsored Research
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3890

Type of business: **Other Educational**

Innovative Claims

The ALVINN-on-a-chip project will build and deploy a single-chip VLSI solution that provides direct high-level sensory-motor control from complex image information. This approach offers many advantages —

- adaptability,
- low latency,
- high frame rate,
- low power,
- compactness,
- ruggedness, and
- low cost.

Previous attempts to build such a device have been limited to servoing on relatively simple (typically single-spot) image input. No previous work offers a trainable, integrated solution that meets the strict real-time requirements of robotic tasks like autonomous road following and the Intelligent Highway Vehicle System (IHVS).

Deliverables Associated with the Proposed Research

At the completion of the proposed research, we will deliver —

- the design for the ALVINN-on-a-chip VLSI computational sensor,
- the algorithm (and its software implementation) used to train the device for autonomous road following tasks,
- plans for a fully-functional driver-assist system based on the VLSI device, and
- experimental results obtained through testing of the system in actual use controlling the HMMWV (High-Mobility Multi-purpose Wheeled Vehicle) robotic vehicle.

We make no proprietary claims.

Cost Estimates, Schedule and Milestones

Year One: $ 450,000

Develop, design and fabricate first-generation ALVINN-on-a-chip.

Year Two: 350,000

Design and construct prototype ALVINN-camera system. Characterize device performance and demonstrate its functionality. Develop training algorithms.

Year Three: 350,000

Train the ALVINN-on-a-chip system. Evaluate the system through actual use controlling the HMMWV robotic vehicle on a public highway.

Total Budget: $1,150,000

Technical Rational and Approach

The ALVINN-on-a-chip project proposes to build and deploy an intelligent VLSI image-motor sensor. This device provides a single-chip link from sensory input to motor output. Circuitry to perform parallel neural computations is integrated with an array of photosensors. Internally, complex, high-level control is performed by an ALVINN (Autonomous Land Vehicle In a Neural Network) neural network, using locally-sensed, complex image information.

A computational sensor intelligently extracts information from transduced data at the point of sensing. The bottleneck between transducer and computer, present in traditional sensing system implementations, is eliminated. Recent work at Carnegie Mellon University (CMU) on a VLSI computational sensor for fast range imaging (done under DARPA contract MIP-9047590) demonstrates the power of this approach. The light-stripe range sensor chip acquires a complete 32×32 pixel frame of range data in a millisecond, a range-image rate of $1,000$ frames/second. This represents a one-hundred fold improvement in frame rate over current range-imaging techniques.

In addition, CMU is a world leader in the application of neural networks for autonomous robotic-vehicle control. The Navigation Laboratory (NAVLAB) work is part of ongoing DARPA-sponsored research projects (contracts DACA76-89-C-0014 and DAAE07-90-C-R059). The ALVINN (Autonomous Land Vehicle In a Neural Network) system steers the robotic NAVLAB vehicle by watching the road. ALVINN has successfully driven on public highways at speeds of 55 miles/hour for distances over 20 miles and holds the world speed-distance record for robotic driving.

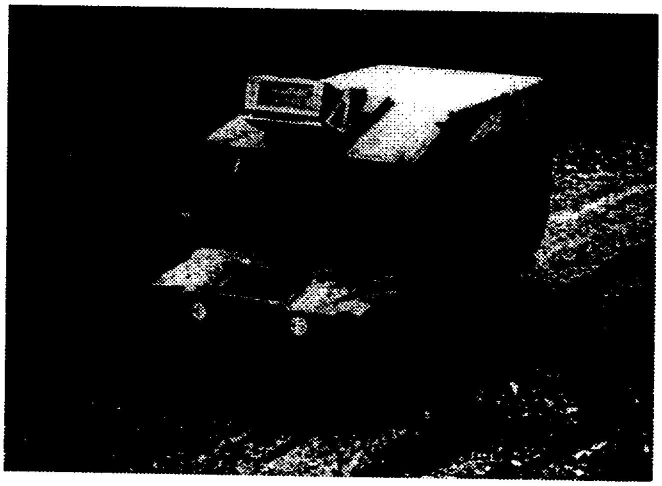

Figure 1: The HMMWV testbed.

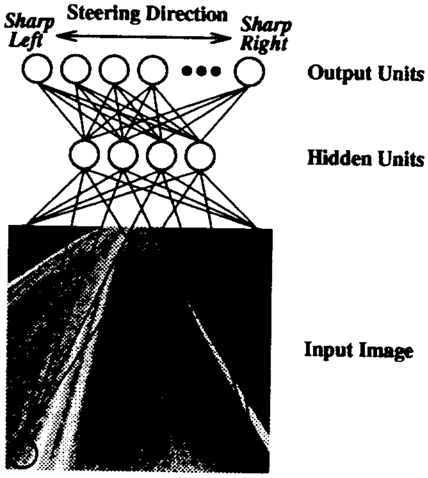

Figure 2: The ALVINN network.

VLSI Computational Sensors

Using VLSI technology, raw sensing and computation are integrated. The result is a new robotic sensor methodology — *computational sensing*. Computational sensors tightly couple sensing and processing on VLSI substrates. Sophisticated imaging tasks are performed by directly operating on sensed information. Previous work on focal-plane processors has been mostly limited to image preprocessing operations. ALVINN-on-a-chip will perform a complex, high-level sensory-motor task.

Latency, the time required to compute an output decision from sensory input, is difficult to reduce when sensing and processing are done separately. Local processing of image information reduces system latency while increasing data throughput. Both are necessary for tasks with strict real-time requirements like automonous road following.

ALVINN: A Neural Network for Road Following

ALVINN is a neural network for driving an autonomous vehicle over single and multi-lane roads. Image information, acquired using a video camera, is processed by the neural network to generate the the vehicle heading control required to stay on the road (Figure2). An implementation of ALVINN running in software on SUN4 workstations has successfully driven the CMU NAVLAB I and NAVLAB II HMMWV at 55 miles/hour speeds for distances exceeding 20 miles.

Neural networks are rapidly programmed for new situations. ALVINN is trained on-the-fly. Road images are processed as the vehicle is driven by a person. Vehicle heading, as steered by the human driver, provides feedback. The network learns to steer as the human driver did when similar driving situations are encountered.

A VLSI Computational Sensor in the Driver's Seat

Image-based control of a vehicle at high speeds is a demanding real-time task. While an image sensor generates vast amounts of data, only a small fraction of the information is relevant. Human drivers use their experience to extract relevant information from what they see. The ALVINN neural network provides a similar capability, extracting the information required to stay on the road. Through a training process, the network learns to filter out image details not relevant to driving.

ALVINN-on-a-chip performs neural processing directly on sensed intensity information. The output bandwidth that must be analyzed by higher levels of processing is greatly reduced and much of the latency between input and decision is eliminated. Our ALVINN-on-a-chip system will achieve rates of 100 frames/second or better, processing the 32 × 32 pixel images it acquires.

An ALVINN Camera

It is not practical to train the ALVINN neural network on real images alone. Human drivers do too well at maintaining vehicle heading and do not provide information necessary to recover from driving errors (where the vehicle might stray off the road, for example). A means of providing synthetic images to the sensor during training is required.

We will build an "ALVINN camera" to facilitate the use of ALVINN-on-a-chip in real world applications. Its essential components, in addition to the computational sensor, include imaging optics, an LCD (Liquid-Crystal Display) panel, backlighting facilities and a shutter. The LCD is used to present precise training images to the sensor, under program control.

Research Impact

Autonomous driving is an important practical problem. Increases in the number of vehicles on our highways at times push human drivers to the limit. Highways will be made safer if systems to assist human drivers can be deployed. Truck convoys will require that only the lead vehicle be driven by a person. The other vehicles can be controlled autonomously by ALVINN systems trained to follow the truck in front. Both of these scenarios, and many others, are made possible using technology developed in this research.

ALVINN-on-a-chip provides general feature recognition capability. Autonomous driving is only one example of a robotic task that can benefit from a neural network which operates directly on sensed images. The ability to rapidly track features is essential to many robotic applications such as manipulation, navigation, and human-computer interfaces. The Self-Mobile Space Manipulator (SM^2), developed by the CMU Robotics Institute, uses an ALVINN network to identify its foot placements, enabling it to "walk" on space-station trusswork. Real-time detection and tracking of attachment points on trusswork nodes for the mobile robot is performed. Eye tracking, the ability to determine the direction of a person's gaze, is another application we have found to be performed well by ALVINN.

Other Research in this Area

There has been a considerable amount of research into computational sensors and VLSI implementation of neural networks. Research issues under investigation include network architecture, analog verses digital implementation and representation of connection strengths on a chip. Some chips are designed for a specific task (Lyon's Optical Mouse from Xerox PARC is an example) while others are programmable and can be adapted to the task at hand. One especially interesting effort, that addresses the permanent weight storage problem, was done at Intel by Holler and Tam. Intel's ETANN (Electrically Trainable Artificial Neural Network) uses floating-gate non-volatile memory elements for storage of analog connection strengths. Weights are changed in a manner similar to that used to write data into standard EEPROM (Electrically-Erasable Programmable Read-Only Memory). The neural architecture of the ETANN chip consists of 8,192 synapses which fully interconnect 64 neurons and their inputs. Each synapse performs signed analog multiplication of its input by the stored weight. The result is summed and then thresholded for output. The chip performs on the order of 10^{10}th low-precision multiplications per second, several orders of magnitude more processing than computer-based neural-network implementations.

Perhaps the best known work on computational imaging sensors is that being done by Carver Mead and his group at Caltech. In the *Silicon Retina*, a model of the computation performed by the first layer of visual processing is integrated. A two-dimensional resistive network is used to perform the spatially-weighted average of photoreceptor inputs. The output at a given node is the sum intensity inputs in its vicinity, each weighted inversely by the distance from the output node position. The follow on *Adaptive Retina* chip uses floating-gate devices to compensate for variations in transistor offset mismatch.

Key Personnel

Takeo Kanade — Year 1: 10% Year 2: 10% Year 3: 20%

Project leader. Will work in the areas of sensor calibration, perception and robotic application.

Andrew Gruss — Year 1: 30% Year 2: 30% Year 3: 20%

An expert in the design and application of VLSI computational sensing. Responsible for mapping the ALVINN architecture onto silicon and for development of the system electronics and optics.

Dean Pomerleau — Year 1: 10% Year 2: 20% Year 3: 40%

Architect of the ALVINN connectionist system for autonomous road following. Responsible for development of training algorithms and deployment of ALVINN-on-a-chip.

L. Richard Carley — Year 1: 20% Year 2: 10% Year 3: 10%

Expert in the area of analog VLSI design.