Fish Inspection System Using a Parallel Neural Network Chip and Image Knowledge Builder Application

Anne Menendez, Guy Paillet

General Vision Inc.

1150, Industrial Avenue Suite A

Petaluma CA 94952 USA

anne@general-vision.com, guy@general-vision.com

Abstract

A generic image learning system, CogniSight®, is being used for the inspection of fishes before filleting offshore. More than thirty systems have been deployed on seven fishing vessels in Norway and Iceland over the past three years. Each CogniSight uses four neural network chips (a total of 312 neurons) based on a natively parallel hardwired architecture performing real time learning and non-linear classification (RBF). These systems are trained by the ship crew using Image Knowledge Builder, a "show and tell" interface for easy training and validation. Fishermen can reinforce the learning at anytime when needed. The use of CogniSight has reduced significantly the number of crewmembers on the boats (by up to six persons) and the time at sea has shortened by 15%. The prompt and strong return of the investment to the fishing fleet has increased significantly the market shares of Pisces Industries, the company integrating CogniSight systems to its filleting machines.

Introduction

The fish industry is very competitive. Fleet owners are very interested in filling their boats as fast as possible with the highest quality and the least personnel on board to reserve maximum occupancy for their refrigerated storage. During an expedition, which can last between 1 to 2 weeks, the fish processing machinery operates 24/7. Typically, fishes are brought on the boat and dropped into metal pockets, which convey them through cleaning, cutting and filleting machines. Anomalies, which must be detected at the beginning of the chain, include a fish of the wrong species or a damaged fish so it is rejected immediately. In addition, the presence of more than one fish in a pocket, or the improper orientation of a single fish must be detected to avoid jamming the cutting or filleting machines. This type of real-time inspection is not easy to deploy with conventional image processing tools since the size, shape and scales of fishes are difficult to model mathematically. Their aspect can also change depending on the location of the expedition as well as the season of the year. Finally and most importantly, the inspection system must be very easy to use for the fishermen since a software programmer has no place on-boat (to fix a software bug or change an image processing algorithm).

CogniSight is a Registered Mark of General Vision Inc.

Figure 1, The Engey (Iceland) is operating seven CogniSight

Several attempts have been done by Pisces Industries to solve this problem with traditional combinations of camera, frame grabber, PC and image processing software. None have lead to a usable offshore system because of the high non-linearity of the problem.

A neural network approach was the only possible way to deliver a system trainable by the fishermen themselves and very adaptive. A hardware neural network was the best way to deliver a system reliable and fast, with a small size and affordable cost. Typical fish species to be recognized include ill-defined herrings or mackerels.

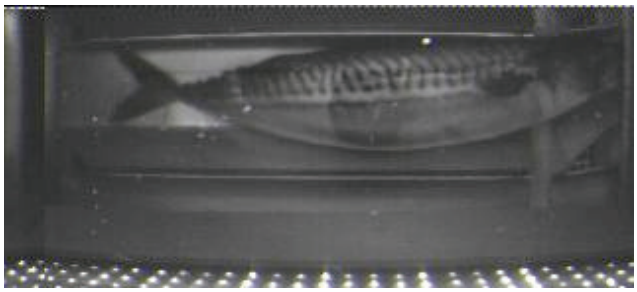

Figure 2, a well positioned mackerel

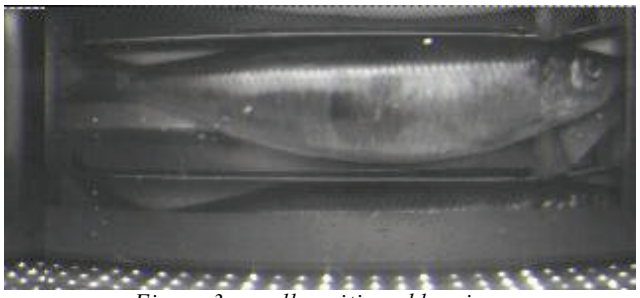

Figure 3, a well positioned herring

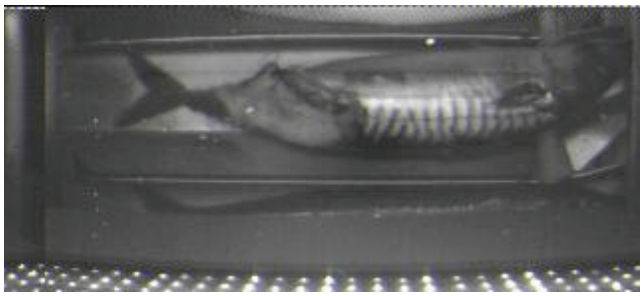

Figure 4, a damaged mackerel to be ejected

Silicon Neural Network justification

Due to the highly variable aspect of a fish, a mathematical modeling was not an option. In addition all the "by catch" which are random species have to be rejected but cannot be learned due to their infinite shapes and textures. In order for the inspection system to operate properly the concept of "Uncertain" and "Unknown" was critical. In addition due to the stringent conditions at sea it was mandatory to provide reliable operation 24/7 within a minimum space and without mechanical components (such as PC fan). Speed was also essential, 360 to 600 fishes per minutes. Usage of statistical analysis on a computer, at such small footprint, high speed and low cost was not a viable solution. In real life, the key is not necessarily to have the absolutely best classifier but instead to have a solution which can solve a problem taking into account its economical aspect as well. In 2003 the ZISC® [1] was the only available neural network chip meeting these constraints. Its RCE [2] (Restricted Coulomb Energy) feed-forward network offers a highly non linear classifier and allows both unknown and uncertainty detection.

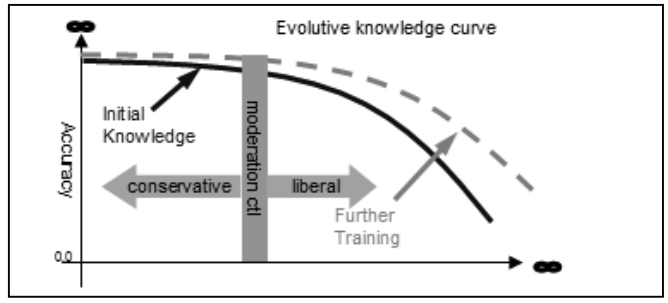

Figure 5 Classification moderation in CogniSight

Learning and Reinforcement

One of the important features was the ability to refine the learning "on board" since a vessel can stay two weeks without getting ashore. An additional aspect was the ability to adjust the "throughput versus accuracy" of the recognition engine by increasing or decreasing the severity on the fish's quality depending upon the situation or season. This is achieved in CogniSight by editing a single parameter: the value of the Maximum Influence Field of the neurons which controls their level of conservatism during the learning process. The higher this value the more liberal is the recognition, the smaller, the more conservatism.

The Solution

A CogniSight system is composed of a vision sensor, a silicon neural network and a recognition engine on FPGA acting as their glue logic, that is extracting a feature vector from each video frame and reading the response of the network after the broadcast of each signature. CogniSight is mounted in a waterproof enclosure above the pocket conveyor just before the filleting machine.

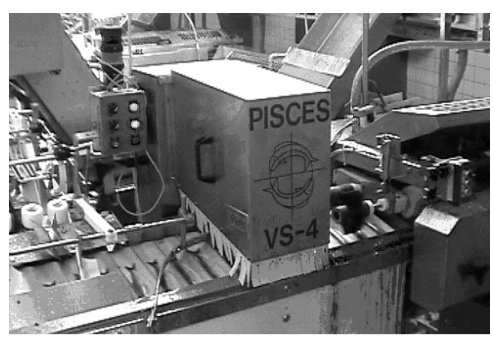

Figure 6. The CogniSight system installed on the filleting line

The autonomy of the recognition relies on the parallel neural network capable of learning by examples and generating models automatically. It can recognize patterns, which are identical or similar to the models stored in the

neurons, and produces three types of responses: (1) a status response which indicates if the classification is identified, uncertain or unknown; (2) a global response which is the category of the first neuron with the best match to an existing model (i.e. with the smallest distance to the input vector), if any; (3) a detailed response which is the category and distance value of all the "firing" neurons read in increasing order of distances. The detailed response of the network can be useful to build hypothesis and leverage uncertainties. In the case of the fish inspection, the global response has proven sufficient because the teaching was done easily but thoroughly on many examples thanks to the Image Knowledge Builder tool delivered with the system. As a result, the system produces few uncertainties and the best match response gives 98% accuracy.

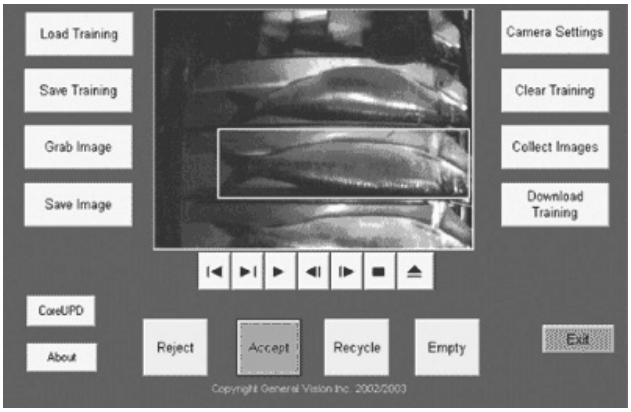

Figure 7, Control panel for on line learning and additional tools

On the filleting line, the number of classifications is limited to the four categories Accept, Reject, Recycle or Empty, but the belonging to a category can include multiple visual criteria. An Accept is a pocket containing a fish of the right species (herring for example), which is not damaged, and in the proper position to enter the filleting machine. A Reject is a fish of the right species but damaged, or a fish of the wrong species. A Recycle is a pocket showing a fish of the right species, not damaged but improperly oriented. A Recycle can also be a pocket with more than one fish of the correct species. The idea of the Recycle category is to eject the fish on a vibrating table so it goes back to the beginning of the line to be dropped into a new pocket. When teaching the system, the fishermen just know to which category a fish belongs. They do not have to worry about describing their rules for decision and simply have to reinforce learning by clicking at the appropriate button on the touch screen panel if they see that the camera is making mistakes.

This ease and freedom of tutoring was the key selling point of the system. The neural network learns the examples and builds the decision space accordingly whether highly non linear or not. The possible drawback of too much freedom in tutoring is that a fish, which looks damaged to a fisherman, can pass as acceptable for another fisherman (or the same fisherman, but on another day). Such contradiction, if any, deteriorates the knowledge by creating "degenerated" neurons. To circumvent this risk, CogniSight is delivered with a software tool for training and validation which allows to teach on many images collected with the sensor and review the consistency of the recognition. This software is called "Image Knowledge Builder™ " and features a very simple and practical user interface to test the accuracy and throughput of the recognition on many images before loading it to the sensor.

Image Knowledge Builder runs a simulation of the silicon neural network. It can save the contents of the neurons into a Image Knowledge File format readable by CogniSight. This file transfer is equivalent to a knowledge transfer and once the silicon neurons are loaded with reference patterns and associated categories, CogniSight can execute the same recognition as Image Knowledge Builder, only at full speed on live images with no connection to a PC. The recognized category is transmitted to output lines indicating if the content of the pocket is an Accept, Reject, Recycle or Empty. This signal is sent to a PLC which itself controls two brushes sending the fishes in a reject or recycle bin when applicable.

CogniSight can be connected to the Ethernet local area network installed on the boat for four types of operations: (1) View the images on screen to adjust the camera settings at the time of installation; (2) Collect images to proceed with the training of the neural network with Image Knowledge Builder; (3) Load an image knowledge file (ikf file) into the neural network of the camera; (4) View statistics accumulated on the camera and reporting the number of acceptable and non-acceptable fishes

Use of AI Technology

Feature Extraction

The feature extraction is executed by an FPGA (Field Programmable Gate Area), which converts the pixel values received from the CCD sensor (640 x 480 pixels) into a feature vector of 256 bytes. It is calculated over a region of interest specified in the video frame which can range from 16x16 pixels to the full frame. This transform is a spatial and grey level integration based on a "best fit" function of 256 blocks within the selected region. The resulting 256 components are transmitted to the neurons and response is read back before the next frame starts.

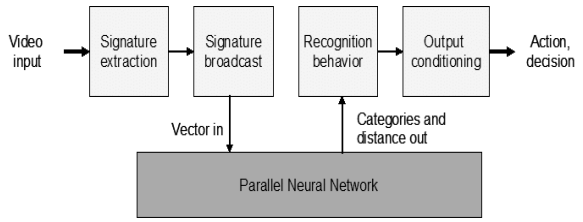

Figure 8, CogniSight image recognition engine architecture

RBF Classifier with Automatic Model Generator

The incoming feature vector is broadcasted to a hardware neural network of 312 neurons. This network comprises four semiconductor chips (ZISC) with 78 neurons per chip connected in parallel. The neurons "react" to the incoming feature vector by evaluating the similarity with their reference vector (stored in their memory during the training). Their network implements a Radial basis Function (RBF) classifier, derived from the compound classifier [3] and the RCE.

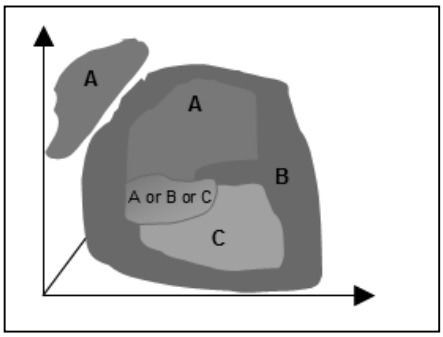

Figure 9, RCE decision space mapping

When committed (or trained), a neuron evaluates if an incident pattern is similar enough to its stored pattern (e.g. model) by calculating an L1 distance. If so, it will output its response to the global response bus. The similarity domain of a neuron is self-mapped during the training process and does not require manual adjustment. If many neurons respond to the same stimuli, a WTA (Winner Takes All) patented scheme allows retrieving in a fully parallel manner the best response(s) of these neurons. The classification can then be determined as positively identified or uncertain. If all the neurons firing with the same smallest distance value are in agreement with the identified category, the classification is said Identified. If, on the contrary, these firing neurons return different category values, the classification is possible but with some level of uncertainty. Uncertainty can be waived using different methods, but in the case of the fish inspection, a conservative training scheme has allowed to obtain good accuracy by simply reading out the response of the first neuron on the list.

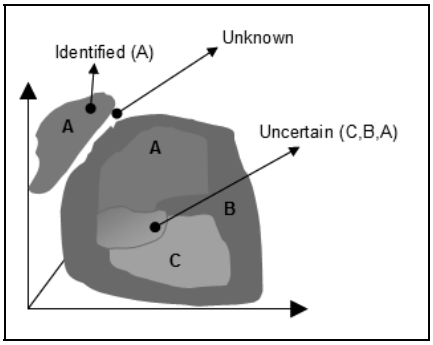

Figure 10, Decisions into feature space

Generation of Portable Knowledge

The neural network embedded in CogniSight is easily trained using the Image Knowledge Builder software. This application features batch utilities to automatically extract feature vectors from annotated images and broadcast these features to the neural network repetitively until the knowledge is stable. Stability is achieved when the learning of the vectors no longer creates any new neuron. Reporting utilities help identify difficult or complex classifications, pinpoint possible erroneous or inconsistent annotations, and evaluate compromises between throughput and accuracy.

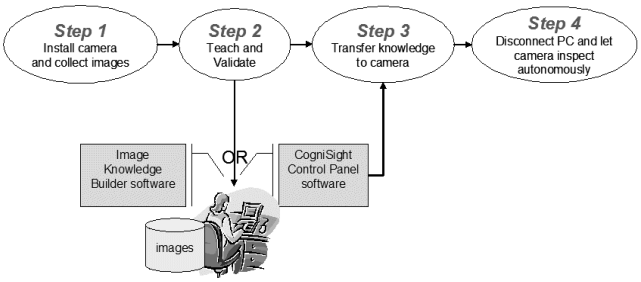

Figure 11, Methodology flow chart

An Image Knowledge File (IKF) can be stored for use on the same boat during a next fishing trip, or it can be transferred for use on other CogniSight systems installed on other boats running the same type of expeditions. Presently Pisces installed systems on fleet fishing for herrings. Their knowledge file has been built over several expeditions taking place throughout the year. The accumulated knowledge can be labeled as "year-around" knowledge. It is composed of less than two hundred models and has already inspected several millions of herrings. In addition, when new conditions are encountered during a fishing campaign the crew can perform reinforcement learning in order to refine the knowledge.

Introducing Silicon Neural Networks (SNN)

Neural network have been extensively discussed since their reappearing in the late eighty's. DARPA[4] recommendation after conducting an extensive survey at that time was to "go silicon". As neural networks are inherently a group of elements having the same basic behavior, they are indeed candidate for massively parallel architecture. While it has been claimed by many publications that neural networks benefit from being parallel, most of the development have been done on standard computers, which suffer three basic limitations:

- They are executing one instruction at a time (sometime four with quad cores).
- A good part of their data bandwidth is dedicated to fetching and decoding their instruction before actually executing them.
- Data is routed through a single bottleneck: the memory bus which in most cases (except for Harvard RISC) also vehicle the instruction. As a result, a neural network

implemented by software on a standard computer cannot be defined as parallel.

Going forward in the massively parallel architecture concept demonstrates that in the case of multiple programmed processors, synchronization between them can become a serious hurdle. The way to overcome these limitations was to design a neuron entity with all the "genetic" material to learn and recall without the need of running program code. In addition this architecture would have to be fully distributed (no supervising unit) and have theoretically unlimited expansion capability. Another constraint was to have a fast and constant learning recognition time, regardless of the number of connected neurons. This was achieved by the ZISC architecture, which was described by the co-author and jointly developed and patented [5] with IBM. The first SNN was the ZISC36 in 1993, followed by the ZISC78. Connections of up to 5,000 neurons (multiple PCI board) were demonstrated. Today the successor of the ZISC, the CogniMem chip (for Cognitive Memory) is being developed and will be available fall 2007.

The Parallel Digital Neuron

The digital neuron consists essentially of a memory storing the learned pattern prototype (or kernel), a hardwired distance evaluation unit, a learning logic and an associative logic. Depending upon its content the neuron has three different states:

- Idle (does not participate)
- Ready to learn (next in line to learn a pattern)
- Committed (learned a pattern associated to a category, and has an influence field)

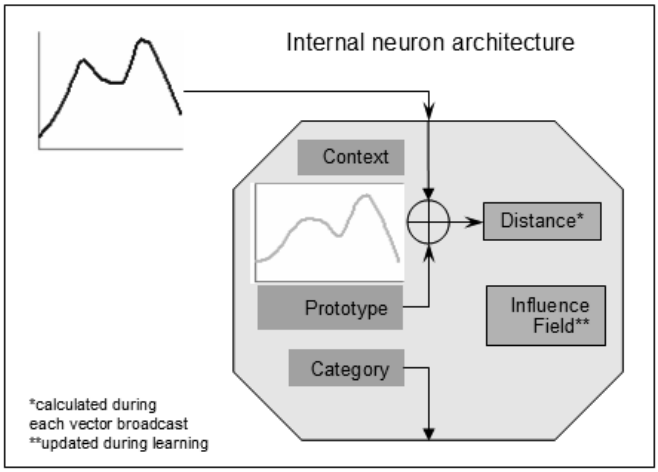

Figure 12, neuron structure

Neuron memory: While the first neurons were using 64 bytes of memory the advance in semiconductor allows now to fit 1024 neurons with 256 bytes each. The information contains into the new 256 bytes memory is indeed richer than previously. This fact actually tends to reduce the number of neurons needed for a given problem. For the inspection of herrings, the original knowledge composed of 300 neurons of 64 bytes has been reduced to 120 neurons of 256 bytes and achieves a similar accuracy. The learned pattern is called a prototype as it is a significant representation of one model of the population.

Distance evaluation unit: The distance evaluation unit computes the L1 distance (accumulation of absolute differences) between the incoming vector (up to 256 components) and the stored pattern. This occurs in parallel for all the committed neurons, at each feeding of a vector component.

Associative logic: The associative logic triggers the output of the category if the evaluated distance falls into the influence field of the neuron.

Learning logic: The learning logic enables a committed neuron to autonomously reduce its influence field (generalization capability) to accommodate the creation of a new neuron if applicable. This is self-adaptation. If no committed neuron identifies the taught category, the Ready-To-Learn (RTL) neuron automatically commits and adjusts its influence field to the distance to its closest neighbor (KNN). All these operations occur inside the neuron and are not under control of an external logic.

Network behavior: While each neuron is fully independent during the learning and the recognition process, all the neurons can "see" the global results. The "search and sort" patented method allows to find the closest distance, "winners takes all", among all the responding neurons in 19 clock cycles (e.g. less than one microsecond) regardless of the number of neurons. This provides the unique ability to learn without the need of program instruction.

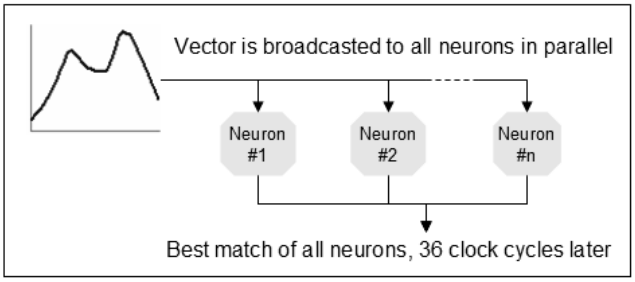

Figure 13, hardware network topology

Use and Payoff

Pisces Industries [6] manufacturer of fish processing equipment has presently installed over 30 systems on five different fleets in Norway, Iceland, Scotland and Denmark. So far, most expeditions have been for herrings and mackerels. The camera inspects at a speed of 360 pockets per minutes on the herring lines, but can go faster. An accuracy of 98% was verified for the classification of 16 tons of fishes with a knowledge based on 80 neurons. The inspection is a tedious job for a human operator, especially considering that he also has to supervise at the same time multiple noisy heavy machinery stations such as feeders, ejectors, vibrating table, the filleting machine, etc. The use of the CogniSight systems for the inspection has contributed to shorten the duration of the expeditions. As a result, a boat can fill its cargo in 5 days instead of 7 days. Fishermen appreciate these shorter expeditions, and sharing bigger catch with fewer co-workers.

Maintenance

Over the past 3 years, no maintenance has been requested. Tuning of the image recognition for different types of expeditions has been handled by the crews. Except for possible electrical failure of the camera, the only serious problem, which can be envisioned, is an insufficient number of neurons if one day an Image Knowledge File requests more than 312 neurons. When this occurs, a CogniMem chip with 1024 neurons should be available and integrated in the next generation of CogniSight systems.

Figure 14, New CogniSight Miniature Trainable Sensor

Conclusion and Perspectives

It has been proven in this application that Artificial Intelligence can provide high return of investment for small entities unrelated to the field. The evolution of the "hardware neural networks" toward a 1000 neurons chip will allow significant cost and size reduction for the present installation. At this time, Artificial Intelligence has become a "commodity" for the Norwegian fishermen. The size and cost reduction of the next CogniSight systems should enable additional control point.

The key to the widespread of these systems is cost reduction (less than US$ 500), the robustness to harsh and changing conditions and the ease of training by the operator of the machine. The hardware neural network allows for a dramatic foot print reduction (close to a matchbox) providing the speed of multiple workstations in parallel. After this first success story, there is a strong possibility that the CogniSight technology associated to Image Knowledge Builder makes a significant contribution to turn Artificial Intelligence into a commodity in many domains related to Vision Machine Learning

References

[1] ZISC (Zero Instruction Set Computer) is a registered mark of IBM Corporation, Armonk NY, USA

[2] Reilly, D. L., Cooper, L. N. and Elbaum, C.: A neural model for category learning, Biological Cybernetics, 45 (1982), 35-41.

[3] Batchelor, B.G. (1974) Practical Approach to Pattern Classification, Plenum, ISBN 0-306-30796-0.

[4] Neural networks study DARPA (1988) AFCEA Press International

[5] US patents numbers 5,621,863; 5,701,397, 5,710,869; 5,717,832; 5,740,326; 6,606,614;
European patents 694854; 694853; 694856; 694855; and equivalent Japan, Canada and Korea.

[6] Pisces Industries, www.pisces-ind.com