Sightech Vision Systems, Inc.

Neuro-RAM

How Hierarchies Facilitate Effective Product Inspection and Decision Making

Self-learning vision, like animal-based vision, employs multi-level hierarchies to arrive at efficient and confident decisions. Here we explain the massive process that reduces complex and changing pixel-based real-world input visual data to a useful output decision. This discussion is *informational only* and describes built-in mechanisms that automatically deliver what the user needs. This mechanisms are precisely what makes our Neuro-RAM vision systems the absolutely easiest to use.

Published 2005

Sightech Vision Systems, Inc.
6580 Via del Oro
San Jose, CA 95126
Tel: 408.282.3770 Fax: 408.413-2600
Email: sales@Sightech.com
Web: www.Sightech.com

Table of Contents:

Basic principal: Animals must, if they see something once, be able to "see" it again. This is not as easy as it would appear. Animal vision is magic.

Visual data received by animals from the real world is, on the pixel level, very complex. Consider for example, that an animal only had 4 8-bit graylevel pixels of vision. At first glance this 4 bytes of image data seems very small and nearly useless. If an animal remembers this data and knows when it "sees it again", it must revisit a member of an image space that has 2 to the 32 power possibilities (2^(4*8))! This is means there is one chance in 4 billion that an animal will randomly see and recognize an image again. If a similar scene is presented, then the chances may increase way up to 1 in 2 million or so. At 30 frames per second, it would take more than 18 hours of gazing at this tiny 4 pixel image before the image of interest is seen again.

Imagine that instead of a 4 pixel image, we are looking at a 307,200 (640x480) graylevel image. You can easily see that this image will never be revisited in the lifetime of the Universe let alone the lifetime of an animal! And, to make things worse, a color image with the same number of pixels has 3 times as many bytes of data.

If the same scene is presented again, to us, it looks to be the same scene as before, but be aware that EVERY PIXEL is different – it fails to match the previously learned image anywhere! The real-world never presents an image the same twice – at the pixel level, it is very different EVERYWHERE in the image.

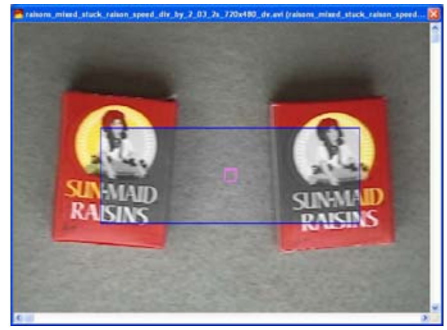
We come to the important question: How can we learn anything and every be able to "see it again" when it is impossible, due to the data complexity and changeability, to get the same data ever again?

Decision hierarchies come to the rescue. By employing deep decision hierarchies, we do get successful matches (we "see it again") at the lowest levels of the hierarchy. By employing a sophisticated voting system that propagates up the levels of the hierarchy, we can successfully get consolidated recognition decisions on VGA (640x480) and larger images at the image level. This is the magic of hierarchies.

Sightech's Neuro-RAM technology automatically uses 6 to 8 levels of decision hierarchy

The levels employed by Sightech Vision System's Neuro-RAM technology are:

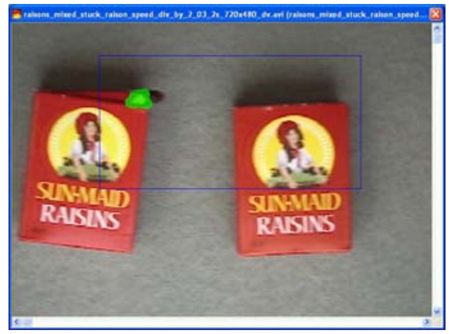
1) Color to graylevel conversion: This level converts, with a fixed formula, 3 color bytes per pixel to 1 byte per pixel. Since there are 65,536 different colors that map to exactly the same 8-bit graylevel value, this simple reduction step is a large one indeed. By reducing the image data size we are increasing the chances of being able to "see it again".



2) Conversion to binary image data: Convolution operators such as "Simple Thresholding", "Sobel" or "Mexican Hat", reduce a graylevel image to a binary black-or-white (BW) image. In the tiny 4 pixel graylevel image we talked about above, these operators would reduce a 32-bit image value to 4-bits. In this extremely small example, 1,048,576 different 4-pixel images map to one 4_bit image. This is an enormous reduction! Some of these conversions to binary images are what call "edge detection".

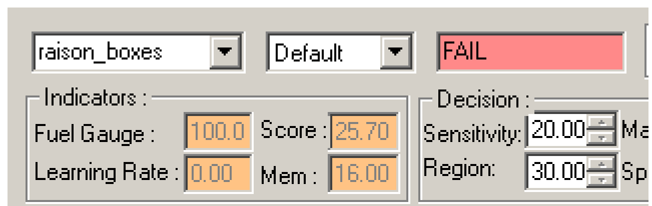
3) Learned, feature-level decision making: This level is the first level using trained conversion. Based on implemented feature constructs, that sometimes include color, a group of binary and/or color pixels are reduced to 1 bit. For example, a feature encompassing 32 pixels of localized information is reduced to 1 bit by this process. For a binary image, the output of level 1 and 2, this is a 4 billion to one reduction. The feature-level decision outputs on bit! In the image we place the feature-level decision bit in the center of the cluster of pixels that define the feature. There is, therefore, a feature-level decision at every pixel position of the image.

4) Regional decisions: This is an important decision consolidation stage which uses a 10 by 10 matrix of feature-based decisions to make a single regional decision. This is a 100 to 1 decision reduction where a value is calculated from the total success of the feature-based decisions within the region. The user can set a regional decision threshold that is compared to the value calculated. The result of this comparison is a binary region-level decision.
Note: The regional decision is placed in the center of the 10x10 features that are involved. Therefore, there is a separate regional decision at every pixel-location in the image data.



5) Multiple-frame decision consolidation: Multiple images may be used to arrive at an even higher level decision by averaging the decision scores over multiple frames. This improves decision confidence even further.

6) Area-level decisions: This decision is highly consolidated and is arrived at by creating a scaled "score" from all the region-level decisions in the defined Area (processing area or ROI). At this level we have arrived at a binary decision, 1 bit of data, that is the final result for the entire image Area.



7) Multi-Area consolidated decisions: The decision outputs from multiple user-defined Areas can be AND'ed, OR'ed, etc. together to create a single camera-level decision.

8) If multiple cameras are involved, the decision process can consolidate at the camera level as well – producing one global and final decision involving different views from different cameras.

Summary of 8 levels in decision hierarchy

In summary, the decision hierarchies are:

1) Color to graylevel conversion using fixed formula
2) Conversion to binary images using fixed operators
3) Trained feature-based decision level
4) Scaled regional decision consolidated from 100 feature-based decisions
5) Scaled score from a moving average calculated from same Area on consecutive image frames.
6) Binary Area-level decision by comparing the score to a settable threshold
7) Multiple Area decision consolidation
8) Multiple camera consolidation

In summary, animals and machines must see the “small picture” in order to see and decide on the “big picture”

Animals and trainable vision obtain high-level decision from a large visual field-of-view. This involves usually from 300k to several million pixels of data. The number of combinations possible in an entire image is beyond galactic in size – this image will never be visited again in the live of the universe. This fact requires the real world and its large number of pixels be observed by animal (or machine) first at a low level. A multitude of low-level input then must undergo severe data reduction and aggregation. Learning processes combined with hierarchies conceptualize and reduce the data further. The result of this large massive process is the high-level detection of an object, or in the case of inspection, an aggregate decision. This results in some action.