CogniSight
Miniature Trainable
Vision Sensor

Hardware Manual

Version 2.0

A product of General Vision Inc.

Table of Contents

Additional documentation
- Micron sensor chip MT9V022 reference guide
- CogniMem reference guide
- CogniSight Technology Guide

Preliminary

Limitation of Liability
General Vision assumes no responsibility or liability for the use of any of this product, conveys no license or title under any patent, copyright, or mask work right to these products, and makes no representations or warranties that these products are free from patent, copyright, or mask work right infringement, unless otherwise specified.

Life support applications — This product is not designed for use in life support appliances, devices, or systems where malfunction of these products can reasonably be expected to result in personal injury. General Vision customers using or selling these products for use in such applications do so at their own risk and agree to fully indemnify General Vision for any damages resulting from such application.

Right to make changes — General Vision reserves the right to make changes, without notice, in the product, including circuits, standard cells, and/or software, described or contained herein in order to improve design and/or performance.

Trademarks
CogniSight is a trademark of General Vision Inc.
Product names listed are trademarks of their respective manufacturers.

Support
support@general-vision.com
Tel 707-765-6150

Specifications

CogniSight MTVS is a miniature trainable vision sensor which can learn and recognize in real-time a region of interest within a video frame. It is surely the first vision sensor with "fully hardwired" image recognition capabilities and sinking less than 150 milliamps...Recognition is performed at 60 frames per second and made possible by loading a knowledge into the neurons of the CogniSight engine.

A knowledge is created by teaching examples of what to recognize. It can be built in real-time or at a deferred time. The category recognized in each frame is transmitted to the output bus and serial bus of the board. It can also be read on demand by a host system via I2C communication.

The knowledge stored in the neurons can be saved. If saved to the Flash memory of the board, the recognition will resume automatically the next time the board is powered up.
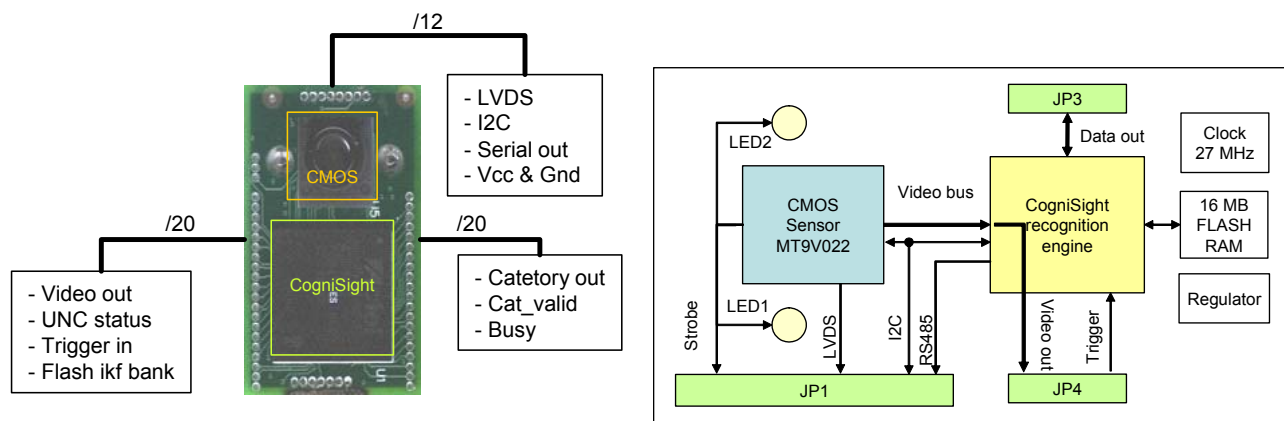
Video acquisition
- Micron CMOS sensor MT9V022
- Monochrome, 752x480 pixel resolution
- Progressive scan, 60 fps
- Global shutter
- External trigger
- Attribute settings via I2C control
- Miniature lens with 4.8 mm focal length
- 2 IR LEDs (870 nm) controlled by the CMOS shutter

CogniSight recognition engine
- Network of 18 neurons hardwired in parallel
- Region of interest (ROI) ranging from 16x16 pixels to full frame
- Recognition of the ROI at frame rate
- Output of the most confident neuron
- Optional filtering of response with uncertainty
- Classification into up to 18 categories

Communication and I/Os
- I2C serial interface at 100-400 Kbits per second
- Serial and parallel response output
- LVDS and parallel video output

Mechanical and electrical specifications
- 26 mm x 40 mm
- 1 ounce
- 3.5v to 12 v DC power supply
- Consumption <<150 mA

Video Acquisition

The sensor attributes such as gain, shutter speed, auto-focus, etc can be changed via I2C control commands. The most common settings are listed below.

Micron sensor slave address= 0xB8

Gain (Addr 0x35)
Shutter (Addr 0x0B)

Auto_Again (Addr 0xAF)
      Auto-Gain is enable if bit 0 is set to 1.

Trigger (Addr 0x07)
      Trigger is enable if bit 3 and 4 are set to 1.

Strobe Out (Addr 0x1B)
      Strobe out is enable if bit 0 is set to 0. The pulse width is equal to the exposure time.

Refer to the Micron manual for the complete description of the sensor registers.

The video signal is available on the LVDS output lines or on the parallel video output bus. Refer to the Micron manual for information on the LVDS signal.

Real-time video recognition engine

The video signal of the CMOS sensor is directly transmitted to the input data bus of the CogniSight chip. As the pixels of a new frame are received by the CogniSight chip, a signature is assembled into a vector of 256 components and transmitted to a recognition logic state machine. The recognition logic broadcast the vector to a neural network and retrieves the response of the neuron with the highest confidence level. If other neurons recognize the input vector with lower confidence levels, they are used to classify the recognition as identified or uncertain. Identified means that all the "firing" neurons recognize the same category. Uncertain means that one or more neurons identify the vector with a different category. A category 0 stands for "unknown" meaning that no neuron has recognized the signature extracted from the video frame.
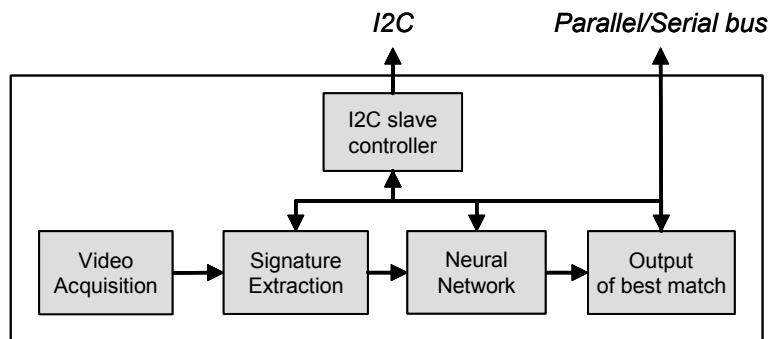
The neurons are taught what to recognize by learning examples. They are capable of generating models automatically as new examples are shown. The contents of the neurons constitutes a knowledge which can be saved. A knowledge can be build "live" on the board or loaded from an existing file or memory.

Setting the recognition engine

The Recognition Status register allows to turn ON or OFF the recognition engine and also to enable the output of the recognized categories to the parallel and serial output bus. The register also reports the status of the recognition in terms of timing and classification status.

CogniSight slave address= 0x94

Recognition Status Register – RSR (Addr 0x1C)

Bit7	Bit6	Bit5	Bit4	Bit3	Bit2	Bit1	Bit0
	Conservative response	Recognition active	Frame valid	Identified recognition status	Uncertain recognition status	Output Enable	Recognition Enable

The RSR can be read and written at any time.
When Recognition Enable is set (Bit0=1), the execution of I2C commands shall be limited to the Read FeatDist and Read FeatCat. The execution of any command should be preceded by an interruption of Recognition Enable (Bit0=0).

Setting the Region of Interest

At each frame, a signature vector is extracted from the pixels located inside a region of interest. Because the neurons can only learn a pattern of 256 bytes, the region of interest must be composed of a maximum of 256 blocks which can be square or rectangular. A block can be as small as 1 pixel if the region of interest occupies less than 256 pixels (for example 16x16 pixels, 64x4 pixels, etc).

Given the size of a region of interest, a common block size can be the smallest square so that 256 of them can fit inside the region. This approach allows to generate a vector data with the maximum amount of information. For some applications, it may happen that it is best to define linear blocks (1 x n, n x 1) or else.

As a result the region of interest is defined by the parameters as described below. They can be accessed in read and write mode.

CogniSight slave address= 0x94

Left (Addr 0x11)
Top (Addr 0x12)
        Position in pixel of the upper left cornet of the region of interest

Block Width (Addr 0x15)
Block Height (Addr 0x16)
        Size in pixels of the blocks constituting the region of interest. These blocks can be square or rectangular. The region of interest can contain up to 256 of them.

Width (Addr 0x13)
        Must be a multiple of Block Width. The sum of Left and Width must be less than the width of the full video frame.

Height (Addr 0x14)

Must be a multiple of Block Height. The sum of Top and Height must be less than the height of the full video frame.

Changing the region of interest takes effect at the next video frame.

Recognition results

When the frame valid signal falls, the recognition engine immediately broadcasts to the neurons the vector which was assembled when frame valid signal was high. The response of the neurons is available within the next 38 clock cycles. The first value read after 19 clock cycles is the smallest distance between the vector and the neurons recognizing it. The second value read 19 clock cycles later is the category value of the first neuron firing with this smallest distance.

CogniSight slave address= 0x94

FeatDist (Addr 0x0D)

Smallest distance between the input vector extracted from the current frame and the pattern stored in the neurons. Read only.

FeatCat (Addr 0x0E)

Category of the first neuron firing with the smallest distance. Read only.

RSR (Addr 0x1C)

This register us updated after each frame and Bit 2 and 3 report if the result of the classification is identified or uncertain. If Bit 2 is equal to 1, the recognized category is not the only one and other neurons (with a same or higher distance) recognize a different category. If Bit 3 is equal to 1, all firing neurons are in agreement with the output category. Depending on the application, the reading of the RSR might be useful to filter or weigh the results.

Category output bus, Cat_valid strobe and UNC status

Category of the first neuron firing with the smallest distance. This value is transmitted to the 8-bit data output bus if Bit 1 of the RSR register is set to 1. In addition, if Bit 6 of the RSR register is set to 1 the output of the category can be filtered out when the classification status is uncertain (RSR Bit 2 =1).

Cat_valid data strobe arises for a duration of two clock cycles when the recognition of a frame is finished. This strobe occurs once during each frame valid low signal.

UNC_I reports if the classification is uncertain or not.

Autonomous power up

CogniSight MTVS features a Flash memory intended to store up to four project files. A project file is composed of the definition of the region of interest, the contents of the neurons and the recognition status register. The selection of the project file is made with the two IKF_ID lines.

At power up of the board, the selected IKF_ID is automatically loaded to the appropriate registers of the CogniSight chip. If the contents of the RSR register is such that bit0=1, recognition automatically resumes.

Neural network access

The neural network embedded in the CogniSight chip features a powerful non-linear classifier and automatic model generator. The recognition engine built into the chip automatically feeds the signatures extracted from the video frames to the network and reads back the best response. As a consequence, direct access to the neural network can be limited to the Teaching command and the Save and Restore operations.

IMPORTANT: Teaching does not have to be executed on live images necessarily. If applicable, it is recommended to use CogniSight MTVS to save images and to build the knowledge of the recognition engine using these images and comprehensive tools such as the Image Knowledge Builder application.

Write Cat (Addr 0x04), value VALUE

Request that the first neuron in the chain which is not already committed learns the signature extracted from the current frame as belonging to the category VALUE.

Save neuron contents

//stop the recognition engine
       Write RSR (Addr 0x1C), value=0x0000
//set the network to save/restore mode
       Write NSR (Addr 0x0D), value=0x0010
//read the content of all the neurons
       For each neuron
       {
              For each vector component (0 to 255), Read COMP (Addr 0x01)
              Read NCR (Addr 0x00)
              Read AIF (Addr 0x05)
              Read CAT (Addr 0x04)
       }
//set the network back to learn/reco mode
       Write NSR (Addr 0x0D), value=0x0010
//restart the recognition engine if applicable
       Write RSR (Addr 0x1C), value=0x0000

Restore neuron contents

//stop the recognition engine
       Write RSR (Addr 0x1C), value= 0x0000
//set the network to save/restore mode
       Write NSR (Addr 0x0D), value=0x0010
//write the content of all the neurons
       For each neuron
       {
              For each vector component (0 to 255), Write COMP (Addr 0x01), value
              Write NCR (Addr 0x00), value
              Write AIF (Addr 0x05), value
              Write CAT (Addr 0x04), value
       }
//set the network back to learn/reco mode
       Write NSR (Addr 0x0D), value=0x0010
//restart the recognition engine if applicable
       Write RSR (Addr 0x1C), value=0x0000

For the complete list of the neural network registers, please refer to the manual "Introduction to the CogniMem Technology".

I2C communication protocol

CogniSight MTVS features an I2C slave controller which can receive Read and Write commands as described below when the Busy signal is low and the RDY signal is high.

I2C Write	I2C Read
Write I2C slave Address + W bit	Write I2C slave Address + W bit
Write 8-bit Register value	Write 8-bit Register value
Write I2C slave Address + W bit	Write I2C slave Address + R bit
Write 16-bit Data value	Read 16-bit Data value

CogniSight slave address   0x94
Micron sensor slave address   0xB8

For more information about the I2C protocol, refer to the Micron sensor chip MT9V022 reference guide.

Connectors

JP4

SAMTEC TMS-112, male header with 12 pins and 0.05" spacing

1	I2C_SCK	I2C clock line. 100 to 400 kbit and 3.3V
2	I2C_SDA	I2C data line. 100 to 400 kbit and 3.3V
3	GND	
4	LVDS+	Low Voltage Differential Signaling used to transmit the video out. 27 MHz base
5	LVDS-	clock and 1.8V +/- 100 mV.
6	VCC In	3.5 to 12 v
7	Reserved	
8	GND	
9	LED OUT	Pulse lasting the duration of the exposure time of the sensor. This value can be set via I2C control bus. This pulse drives the two IR LEDs (870 nm) mounted on the board.
10	RS485+	Low voltage 3.3V serial out used to transmit the recognized category after each
11	RS485-	frame. 115200 bps, 8 bits, no parity, 1 start bit (0), 1 stop bit (1), no echo.
12	GND	

JP1

SAMTEC TMS-120, male header with 20 pins and 0.05" spacing

| 1	VCC In	3.5 to 12 v
2-17	DATA00-15	Category value of the signature extracted from the last video frame.
18	CAT_VALID	Pulse signal indicating that a new category data has been latched on the DATA bus. The pulse lasts one clock cycle and occurs within 10 microseconds of the fall of the frame valid signal.
19	BSY	High when the chip is busy executing a command. Must be low to execute an I2C command received from a master I2C controller
20	GND	

JP3

SAMTEC TMS-120, male header with 20 pins and 0.05" spacing

1	VCC In	3.5 to 12 v
2	VCC IO	3.3v
3	VCC	1.5 v
4	TRIGGER IN	High for continuous grab, Low for stop. Positive pulse for snapshot.
5	Reserved	
6-13	V_DATA	Video data out
14	V_CLK	Video clock signal
15	V_FVALID	Frame valid signal
16	V_LVALID	Line valid signal
17	UNC_L	Classification uncertain if UNC_L=0
18-19	IKF_ID	2-bit identifier of the project file to load at the power up (default=0)
20		GND

Operation summary

The following diagram illustrates the possible operations of CogniSight MTVS.
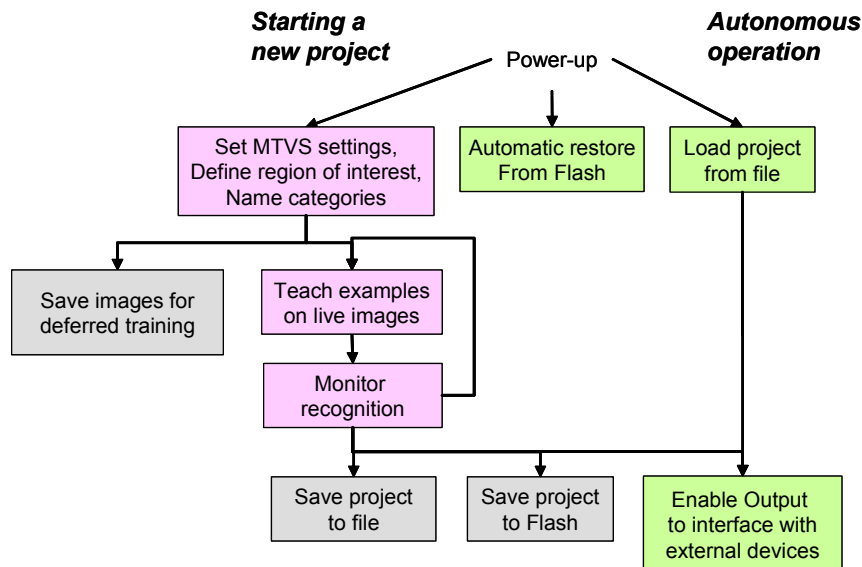

The I2C slave controller of CogniSight MTVS allows to receive and transmit commands between the CogniSight MTVS board and an external controller (PC host, pic controller, microprocessor and else). The steps to start a new application are described below:

- Display the live video
  o Power up the CogniSight MTVS board. The video signal is available on the LVDS output lines or on the parallel video output bus.
  o Once you have access to the display of the video images, adjust the lens to get the proper field of view and focus
  o The sensor attributes such as gain, shutter speed, auto-focus, etc can be changes via I2C control commands.
- Method 1: Quick evaluation of an application
  o Set the coordinates of the region of interest (ROI). The default ROI is a rectangle centered in the video frame and occupying 50% of the field of view.
  o Teach examples seen in the live video by writing the Cat register with the appropriate category value (1 to 5; use 0 to teach background examples)
  o Monitor the recognition by reading the FeatCat register.
  o Teach as necessary until results are accurate, OR proceed with Method 2 as we recommend.
  o If satisfied with the accuracy of the results, you can enable the data output bus by writing 0x03 or 0x63 to the RSR register.
- Method 2: Building and validation of a robust recognition engine
  o Requires a USB adaptor and Image Knowledge Builder
  o Collect many images to proceed with the training and validation of a robust recognition engine using the Image Knowledge Builder (IKB) application.
  o Export the Image Knowledge File to the Flash memory of CogniSight MTVS
  o Reboot CogniSight MTVS and enable the recognition by writing 0x01, 0x61, 0x03 or 0x63 to the RSR register.

Hardware application notes

Depending on your application, you can interface one or multiple CogniSight MTVS with a variety of host controllers.
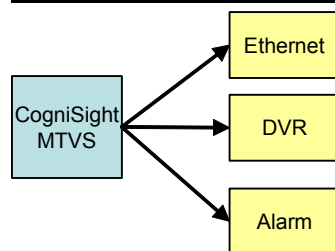
Simple industrial inspection
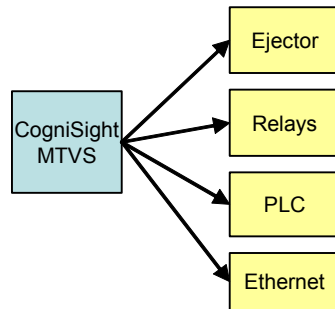
Single CogniSight MTVS, single knowledge for a product line, interface to a variety of industrial I/Os.

On a factory floor, CogniSight MTVS can be interfaced with an ejector to reject or divert parts, with a PC station to generate statistical report about the production and, if needed, record the images showing parts with certain types of defect.

Simple video surveillance

Single CogniSight MTVS, single knowledge per type of objects or scene to detect, interface to a variety of media for selective transmission and recording.

On a surveillance station, CogniSight MTVS can trigger a warning signal, record the date and time a special event occurs and transmit each image showing a particular object over the Internet

Complex part or scene inspection
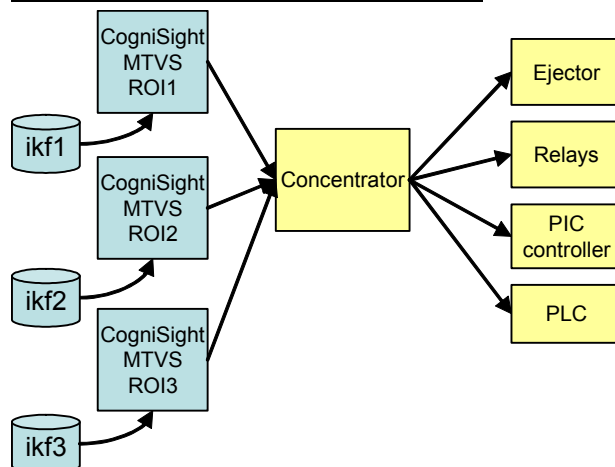
On a factory floor, CogniSight MTVS can be interfaced with an ejector to reject or divert parts, with a PC station to generate statistical report about the production and, if needed, record the images of parts with a certain type of defect.

Multiple CogniSight MTVS with different knowledge files, a result concentrator to generate a global response prior to interfacing to a variety of industrial I/Os.

Target tracking
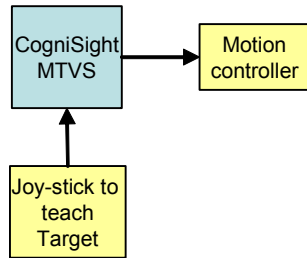
Single CogniSight MTVS interfaced to a motion controller.

For simple target tracking (not identification), the categories reflect the alignment of the target in the region of interest such as centered, left, right, top, bottom. The categories are transmitted to a motion controller which moves the camera in the opposite direction to re-center the target in the ROI.

You may define more categories with N degree angle increments.

Recognition of an object showing up at an unknown distance
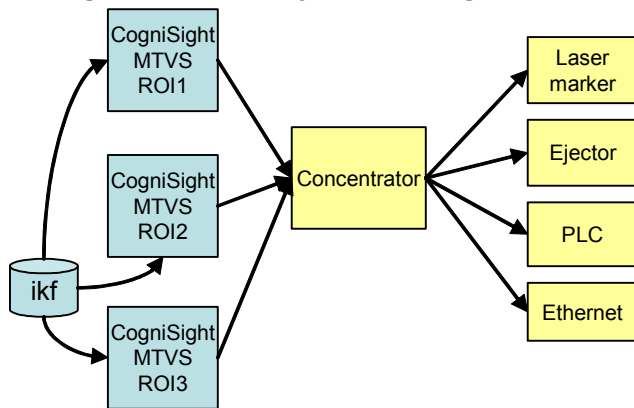
Multiple CogniSight MTVS adjusted with different focal planes (f3>f2>f1) using different ROI (ROI3<ROI2<ROI1), same knowledge file trained at a mid-range focal length, a result concentrator looking for the MTVS which will recognize the object with the highest confidence level when the latter enters the filed of view.
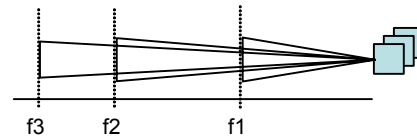
Surface inspection
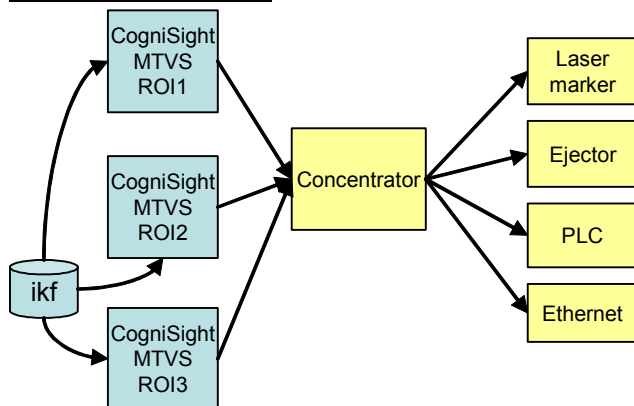
Multiple CogniSight MTVS installed to cover a continuous field of view at a same focal plane, same knowledge file trained to recognize a good texture, a result concentrator programmed to generate an "anomaly" signal if at least one MTVS does not recognize a good texture.

The regions of interest of the multiple MTVS must have the same size and overlap slightly.