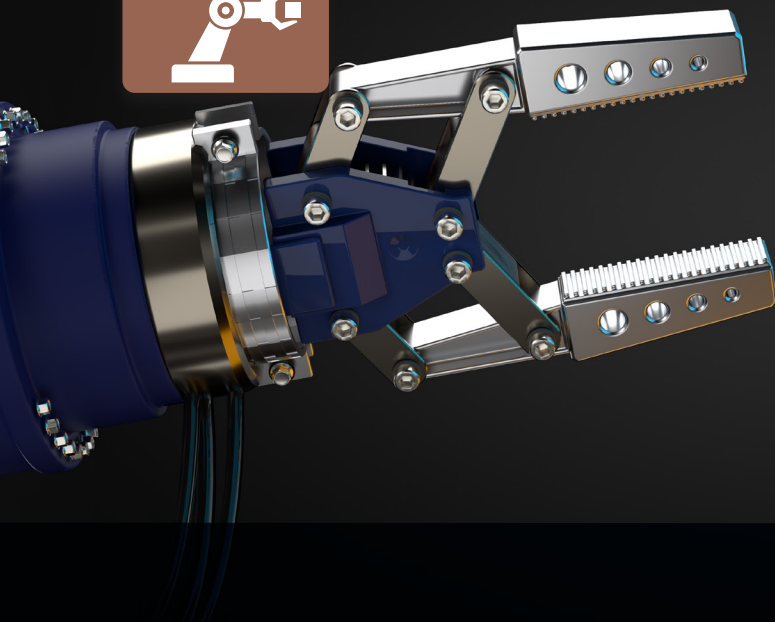




Whitepaper

Industrial



Improved man-machine interfaces using MV and AI with Omron's Human Vision Component - HVC-P2

By Ros Kruger, Director, Supplier Marketing, TTI Europe

Executive Summary

There is no doubt that we have accepted technology into almost every aspect of our lives. One reason for this is perhaps because it operates transparently in the background. The way we abstract ourselves away from technology allows more people to feel its benefits; it isn't necessary to understand how the Internet works in order to use an online shopping service, for example. That wasn't always the case, but technology has a way of changing to meet expectations, particularly when there is a commercial incentive. The recent addition of artificial intelligence (AI) and voice interfaces to create virtual digital assistants is opening up online services to a wider range of users.

The latest developments in machine vision (MV) and AI are also going to create many new commercial opportunities for technology, as they increase the level of abstraction and simplify our interactions with technology so we are hardly aware it is there. We are already able to interact with technology using natural language, the addition of MV and AI will improve the user experience, as will the addition of gesture recognition and the ability to read facial expressions. This will help enable a new generation of more human-focused consumer experiences, as well as provide a significant boost to the development and acceptance of collaborative robots, or cobots. This will fundamentally change the shape of many areas where technology is making a difference, from caring for the sick and elderly to factory automation.

This Whitepaper will explore the area of machine vision as it applies to user interfaces within the area of industrial/commercial cobots, light industrial applications (vending machines, kiosks) and factory automation, with specific focus on the HVC-P2 sensor.

Introduction

There have been enormous advances recently in the field of man-machine interactions, with the use of AI and speech recognition. However, there are still challenges in making interactive systems react more appropriately to the user. One part of the solution will be the use of advanced MV and AI to give the system more information about the person they are interacting with, and the ability to respond appropriately.

The Market

Although there is already a large market for MV, this is dominated by routine inspection and sensing tasks. Advanced systems are becoming available that support more intelligent image analysis. This is expected to be an increasingly important part of the market in future.

The global market for MV systems is estimated to be nearly \$10 billion. It is forecast to grow to about \$15 billion by 2022 (with a CAGR of around 8%).

Hardware makes up the larger part of this, mainly cameras and sensors. The software side is very fragmented and application specific. However, this part of the market is expected to grow significantly as techniques such as machine learning and AI are applied to produce more advanced capabilities that enable new applications.

Growth is driven by several factors including the increasing adoption of Industrial IoT technology which uses vision to improve the productivity of robotic systems and to gather information. In the non-industrial area, novel applications are emerging such as driverless cars, intelligent retail systems, security and access control.

Other factors that will contribute to future growth are the increasing capabilities of the technology and the falling cost of both sensors and compute capability. Also, increased ease of use will make it easier for companies to integrate high-level functions into their products.

Asia-Pacific and Europe make up about 77% of the total market. The Asian market is expected to see high growth because of the focus on the automotive and consumer electronics markets. There is strong competition among consumer electronics companies in the region, which will help drive the greater adoption of MV systems in a range of applications.

Europe is a driving force for machine vision. There is a high level of industrial automation, particularly in the automotive industries in Germany, Italy and France. This is supported by large scale adoption of all types of MV products, from the simplest to the most advanced. There is also a growing market outside of the industrial sector, for example in healthcare. This has created a market that is larger and more diverse than any other.

As well as users of machine vision, there are also many European producers of MV products including sensors, systems and software. Much of this technology is sold to Asia and so Europe benefits from the growth in the Asian market. As a result, the European market is expected to grow by about 7% per year to 2026.

Technology

Current MV systems are based on the work done over many decades in multiple areas including optics, mathematics, physics, computer architecture and software. The technologies that have contributed most are solid-state image sensors, embedded processors and image processing algorithms. This has enabled the development of systems that provide more analysis and "understanding" of what the camera sees.

The development of charge-coupled device (CCD) image sensors goes back to the early 1970s. The concept of the CMOS active pixel image sensor was developed around the same time but semiconductor manufacturing technology did not make mass production practical until the mid 90s. There are currently a handful of manufacturers of CCD and CMOS image sensors, but their availability and low cost has spawned hundreds of camera companies worldwide, several focused on the machine vision and image processing markets.

The use of standardized high-speed, low-cost serial interfaces allows large amounts of image data to be transferred quickly from the camera to the processing system. Some cameras exploit widely available interfaces such as Ethernet, Firewire or USB. This simplifies the task of integrating these into a system. There are also interfaces specifically designed for the needs of cameras, such as CoaXPress (CXP) and Camera Link HS (CLHS), which are used in more specialized applications.

Higher data transfer rates, combined with increased compute performance, enables high resolution images and video to be processed in real time.

Many of the algorithms in use today were developed in the 1960s. This includes techniques for image enhancement, feature extraction and object identification. The functions available have since been extended to include identifying faces, specific individuals and even expressions. Today, the focus is on machine learning and AI to extend the functionality even further to meet the needs of emerging applications. This is leading to the development of smart cameras that do much of the image processing and object recognition internally.

Applications

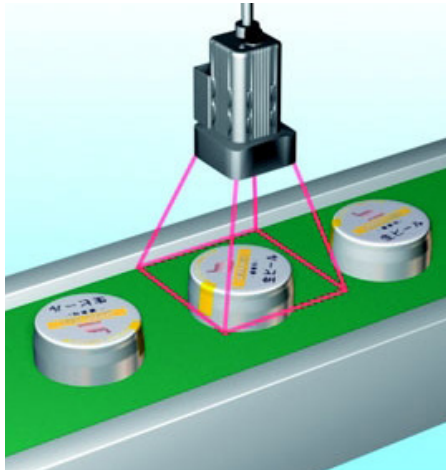


Figure 1: Visual sorting system (source Omron)

The largest application areas, currently, are automated inspection, sorting and quality control. The earliest example of this goes back to the 1930s, with a machine made by Electronic Sorting Machines that used cameras with color filters to identify and select items that were within an acceptable range. This was applied to sorting various types of peas, beans and nuts.

Modern sorting and inspection systems exceed humans for speed and accuracy in many routine tasks. However, as capabilities have improved, a wider range of applications is able to take advantage of MV technologies. This includes:

- Vision-guided robotic systems across the pharmaceutical & chemical, food & beverage, manufacturing and packaging industries
- Agricultural tasks such as planting, tending and harvesting crops
- Automated sorting machines and meat-cutting systems
- Collaborative robots which operate in a joint work-flow with humans
- Advanced driver assistance systems (ADAS) and autonomous vehicles
- Interactive advertising and retail systems
- Security systems for identification and access control

Machine vision now needs to make a step change in capability to support emerging applications and enable new ones, particularly outside the industrial area. This will use machine learning and AI to provide higher level information about the contents of a scene; for example identifying people and what they are doing.

The use of these new technologies will require engineers with new skills. There will be a need for education and training of existing teams. This can be offset by MV products that include the required high-level functionality.

One of the challenges is providing the processing power needed to implement the more advanced algorithms. Many current AI systems rely on being able to offload processing to the cloud. For some embedded systems, this will not be an option, either because the latency is too great or they are not online. These will require sufficient processing in the vision system itself.

Some of the key requirements for adoption of MV systems are:

- **High level function:** for example identifying objects and people
- **Ease of integration:** hardware must be compact, flexible and use standard interfaces
- **Ease of use:** the system designers should not need to understand or implement advanced algorithms to use the system

The Omron HVC-P2 sensor

The Omron Human Vision Components (HVC) sensor is an image recognition system that addresses many of the issues described above. It can not just identify that there are people in the image but also provide detailed information about their facial expression, direction of gaze, gender, age and other information.

The latest device, the HVC-P2, is 10 times faster than the previous product. It is able to track a human body four times a second.

The new hardware consists of an image sensor attached to the main board, which contains the image processing hardware, via a ribbon cable. This provides flexibility for integrating the HVC-P2 into a system. The interface to the HVC-P2 is via a standard UART or USB interface.

The system uses Omron's proprietary OKAO™ Vision software, a proven set of image recognition algorithms developed over 20 years and used in over 500 million digital cameras, mobile phones and surveillance robots around the world. It is able to determine the sizes and position coordinates of human bodies, faces or hands that are detected in the camera image.

The algorithms in OKAO Vision can analyze the facial data gathered from the images to provide information about where each person is looking and estimate information about them. The system can handle up to 35 people in each image.

The system is able to recognize 10 different attributes of people in the image. These are:

1. Face detection
2. Body detection
3. Hand detection
4. Face direction estimation in three directions (pitch, roll, yaw)
5. Gaze estimation in two directions (pitch and yaw)
6. Blink estimation
7. Age estimation
8. Gender estimation
9. The system can recognize five facial expressions: neutral, happiness, surprise, anger and sadness
10. Facial recognition by matching faces against facial data stored in memory

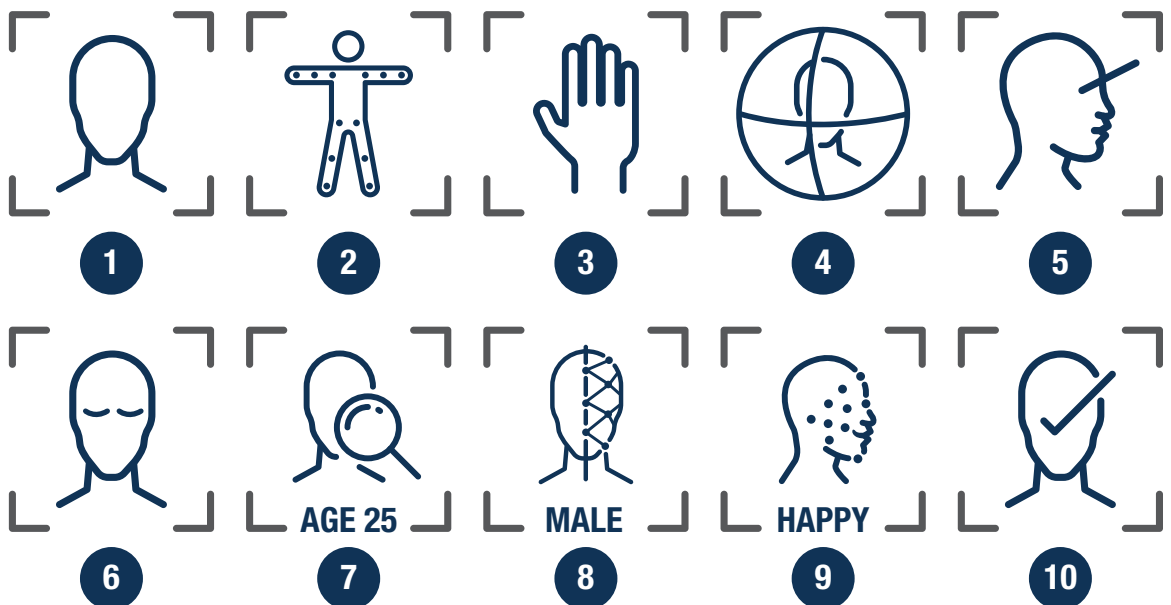


Figure 3: HVC-P2 functions (source Omron)

The image sensing ability is based on data gathered from more than one million faces.

For each attribute the software provides an estimated value, such as the age, the direction the person is looking (as two or three angles) or the facial expression. It also provides a measure of the confidence level of that estimate, from 0 to 1,000 (a higher value indicating a greater confidence).

There are two options for the camera, either a wide angle (90°) or long distance (50°) lens. The choice depends on the detection distance, field of view or other needs of the application. In both cases, the camera has a resolution of 1600 x 1200 pixels.

The long-distance version of the HVC-P2 can detect people at up to 17 meters and estimate the facial attributes at a maximum distance of 3 meters. This could be used, for example, to measure the degree of attention people pay to digital signage in a railway station.

The wide-angle version is capable of covering an area 100 cm by 75 cm from a distance of 50 cm. This is a suitable for monitoring people buying from a vending machine, collecting data to be used later to optimize product refill, product development and marketing activities.

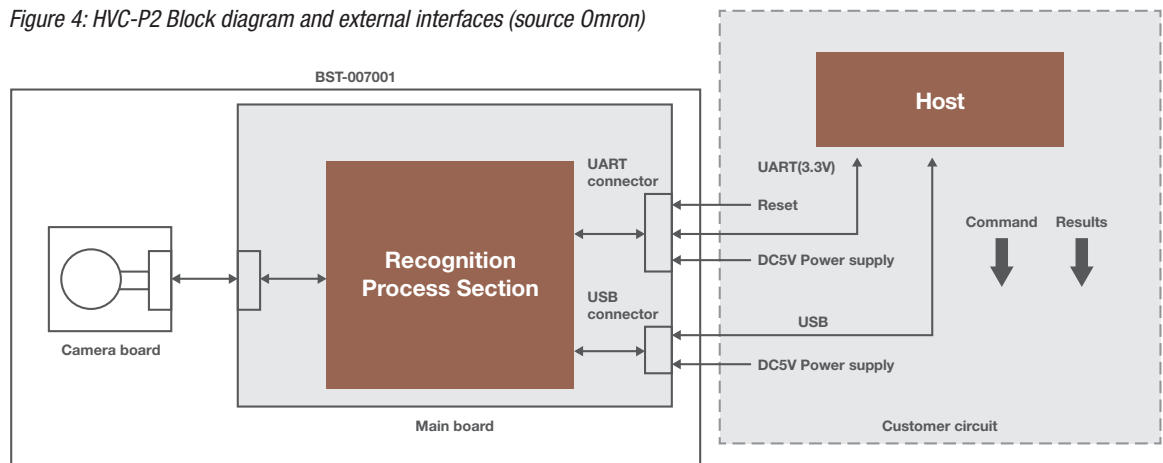
HVC-P2 can be used to develop embedded systems that are aware of, and react appropriately to, people nearby. This can be used in various ways, such as understanding how people react to information presented to them.

Design support

The HVC-P2 is supported with source code and drivers on Linux, Windows and Android systems. There is an evaluation system available for Windows which will let you explore the capabilities of the system interactively. To get you started with your application integration, sample code is provided in C, Python and, for Android, Java.

To help with hardware integration, detailed CAD drawings are available for the main board and camera assembly.

Figure 4: HVC-P2 Block diagram and external interfaces (source Omron)



Target applications

This low cost technology can be embedded in a wide range of systems to add people-sensing to products and services.

There are many non-safety critical applications that can make use of this advanced capability. This includes customer analysis, interactive games, safety systems and many others.

Advertising and vending machines can record information about what people are looking at and for how long. This can provide feedback about how effective signs are, which products they consider before buying, which demographics are interested in each product type, and so on. This enables the information or products offered to be optimized.

Interactive displays can use information from the sensor to adjust the position and size of information to suit the viewer. The content can also be adjusted based on the estimated age, gender or mood of the viewer.



Figure 5: Interactive signs and vending machines
(source Omron)

Security systems can use the HVC-P2 to monitor people entering a building. This can be used to manage elevators based on how many people are awaiting, for example, as well as using facial recognition to control access. Retail stores can count the number of visitors and can use the estimates of age and gender to gather information about their customers.

Monitoring the number of people and their locations can also be used to manage heating, lighting and air conditioning in a building more efficiently. Knowing where people are, the environment can be kept comfortable without wasting energy.

Another possible use is watching out for people in a manufacturing environment by stopping machinery if anyone enters a working area. This is particularly important for people working with collaborative robots (or "cobots") where a robot and a human operator work together on a task in a shared workspace, taking advantage of their specific skills and advantages: the robots provide strength, speed and precision, while the human partners bring intelligence and adaptability. To ensure that the work is carried out safely, it is essential for the robot to know the location of the human worker. Being aware of whether they are paying attention to the work or looking elsewhere can increase safety and efficiency.

ISO standards define the safety requirements for collaborative robots, including the tools they use, objects they handle and the environment they work in. Intelligent MV systems can play a part in ensuring that these requirements are met.

In future, robots may also play a larger role in caring for the elderly or infirm. It will be valuable for them to be able to identify individuals and modify their reaction based on the person's emotions.

Vehicles with various type of driver assistance and autonomy are currently being developed. One important function is to monitor the driver and make sure that they are alert. Tiredness or lapses in attention can be detected and a warning sounded. The ability to identify hands in the image opens the way to the use of gesture control of the infotainment system, for example.

Conclusion

Machine vision has been in use for nearly a century. The applications so far have been dominated by routine industrial inspection and sorting tasks. Improvements in sensors, communications and processing have made these faster and more accurate.

Advances in machine learning and AI are now enabling the development of new human-centric applications that use features like facial recognition and mood detection. These will be valuable in many areas from industrial automation, retail, automotive and the care sector.

Adoption of these technologies will be enabled by products that provide all the functionality required in a way that makes it easy to integrate and easy to use. The Omron HVC-P2 is an example of such a product.



About TTI

TTI, Inc. is the world's leading authorized distributor specialist offering passive, connector, electromechanical, discrete, power supplies and sensor components. TTI's extensive product line and supply chain solutions have made the company the distributor of choice for industrial, defense, aerospace and consumer electronic manufacturers worldwide.

TTI's extensive product line includes: resistors, capacitors, connectors, discretes, potentiometers, trimmers, magnetic and circuit protection components, wire and cable, wire management, identification products, application tools, power supplies, sensors and electromechanical devices. These products are distributed from a broad line of leading manufacturers. TTI strives to be the industry's preferred information source by offering the latest IP&E technology and market information through the online MarketEye Research Center. MarketEye includes articles, technical seminars, RoHS, seminars, industry research reports and much more.

TTI's products, personalized service and custom supply chain solutions have earned us the most preferred passives distributor title by CMP Publications. TTI employs more than 5,600 people at more than 100 locations throughout North America, Europe and Asia.

European Headquarters:

TTI, Inc.
Ganghoferstr. 34
82216 Maisach-Gernlinden
Germany
Tel.: +49 (0)8142 6680 – 0
Fax: +49 (0)8142 6680 – 490
Email: sales@de.ttiinc.com
www.ttieurope.com

Copyright © TTI, Inc. All Rights Reserved.