Optimising machine vision systems in challenging environments

The following white paper outlines recent advances in machine vision-system capabilities and is applicable to those developing systems in the ITS, manufacturing, agriculture, pharmaceutical and similar sectors.

The white paper looks at the design implications for such systems, examining:

- **Overall system design**
- **Lighting and artefact identification**
- **Image processing – on the camera or on the PC**
- **Data management and major transmission standards**
- **System synchronisation and the rise of low-cost methods**

SONY®

1. Introduction and Challenging environments

Balancing the trade-offs of speed and accuracy

The ability to deliver high quality images at high speed has played a crucial role in enabling ever faster throughput. The ability for machine vision systems – involving multiple cameras – to cope with varying light conditions and to trigger at the right point has been vital for countless industries.

The challenges faced vary each time, for example on the controlled environment of the factory floor, systems are optimised for out-and-out speed. In ITS and outdoor applications, systems need to cope with highly variable, and fast changing light. And in agriculture it's not just a high-throughput level that systems need to cope with, but the ability to work with imperfect flows, while still detecting and analysing what's on/ underneath the skin of produce moving down the conveyor belt.

All too often this means perfect synchronisation between the cameras and the various system components is essential.

2. Holistic design

$$\textit{Takt time (T)} = \frac{\textit{Net time available to work (Ta)}}{\textit{Demand (D)}}$$

Machine vision has played an unarguable role in improving both the quality of goods and their production speed, allowing for ever-better error detection and component placement. This is true in industries as diverse as manufacturing, pharmaceuticals, agriculture, ITS, and electronic board assembly. And not just for a subset of processes, but for a near unending set – from car inspection systems that drive down recalls to food sorting and even luxury-watch manufacturing, where previously only the reliance on master craftsmen was trusted.

But, as the technologies improve, emphasis is placed not just on accuracy, but on cost and speed too – without false positives being sent, and without slowing the production line.

Minimising takt plays an increasingly key role in the overall throughput of a system and therefore is essential in cutting the manufacturing costs per unit.

There is no one way to cut takt, instead the system as a whole should be considered. The image sensor's ability to capture frames quickly is an obvious factor, but far from the only one; lighting, the image processing algorithm used, and the module's dynamic range should also be among the methods deployed.

A key shift in this is the move from CCD to CMOS sensors. However, standard CMOS sensors cannot be used for capturing fast-moving objects.

Typically in CMOS sensors, frame rate is maximised by having each row begin the next frame's exposure after completing a read-out. With the time frame being shifted, slight blurs are visible on high-speed images, with the row at the bottom of the sensor almost a whole frame behind the top row.

2. Holistic design

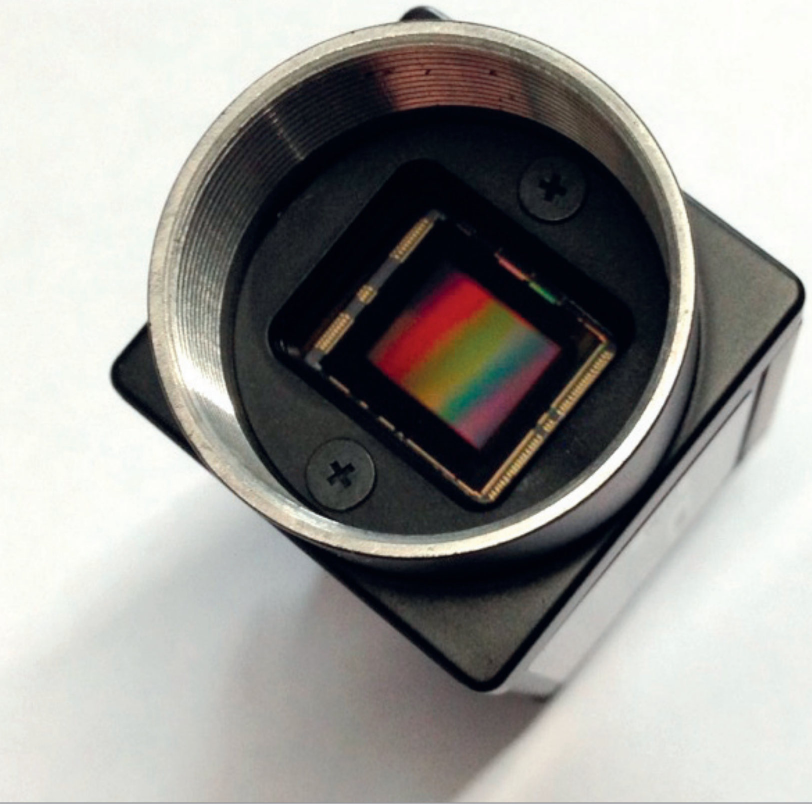
>

This is solved with a global shutter architecture, now used in virtually all CMOS imagers designed for machine vision applications. GSCMOS blanks / activates all pixels at the same time. This prevents the shifting that can result from a conventional CMOS rolling shutter design.

Sony's Pregius global shutter CMOS sensors – such as the IMX264 5.1 MP sensor – are used extensively throughout the machine vision industry and are the gold standard.

But minimising takt and getting the best image for analysis is about more than having a good sensor. You need each component – from the camera sensor, to the system lighting, to the robot arm, to the PC – to be in perfect synchrony and optimised to work together.

Sony has over 30 years of camera design expertise and has the industry's best knowledge on extracting the highest possible image quality from any given Sony sensor.



3. Lighting issues / artefacts

The move to global-shutter CMOS sensors brings several benefits, not least when it comes to lighting.

Moving beyond the camera, lighting is often a challenge in machine vision applications. Intelligent transport systems, for example, have to work with pitch-black nights on the one hand and the bright midday sun at the other extreme, but also has to cope with shadows cast by larger vehicles, headlights, adverse weather conditions, and fast changing environmental conditions as the sun goes in and out of the clouds.

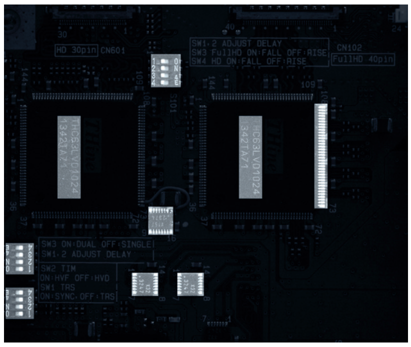
Even in factory set-ups, where lighting can be more easily controlled, it can still represent a challenge. In some applications the demand on throughput means objects have to be captured at exceptionally high speeds without blur and without over or under exposing different areas on the image.

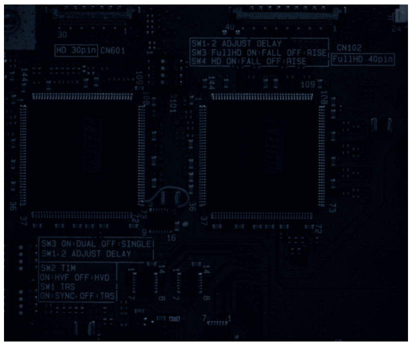
In such applications where speed is vital, exposure time was traditionally cut by setting the captured area's lighting at a high level, or increasing the light sensitivity of the image sensor itself. But this often leads to glare in another part of the image – especially when objects aren't flat or a component is in the shadow of another.

Image processing can be used to partially correct problems, but they also increase computational load, slowing the system and making it harder to distinguish real defects from illumination artefacts.

Global shutter CMOS sensors, however, add another way to manage this exposure, controlling the capture at the pixel level, and through the use of high frame rates – achieved through using CMOS imagers – it is possible to overcome these problems of illumination consistency.

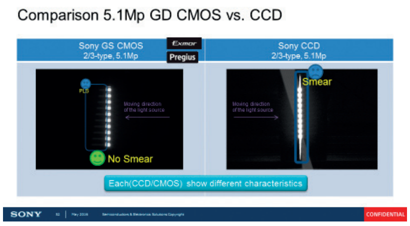
Controlling light at high speed means the camera and the lighting system need to communicate and be in sync, with each camera / lighting system firing precisely. This has traditionally demanded dedicated and costly hardware such as GPS chips, which are accurate to a few nanoseconds. But accurate, system-level synchronisation is now possible with GigE cameras via the use of the low-cost, software-based precision time protocol IEEE1588 – see section 6. Synchronisation.





XCG-SG510 with and without AreaGain

4. Image processing



Dynamic range and multiple captures

Wide Dynamic Range (WDR), where multiple shots are taken in quick sequence with differing exposure times, is used to enhance dark sections and reduce glare on bright sections and thereby optimise the overall image. The resulting composite image has a far higher bit depth than a single image and enables a system to capture areas that would otherwise be lost through glare or shadowing.

These composite images will also increase an image's overall sharpness and counteract effects such as heat shimmer, blurring caused by long-exposure times on moving objects; or shifts in position from product to product on subsequent short exposures.

Reference frames

Modern CMOS image sensors, such as Sony's Pregius IMX264, are capable of altering the effective brightness of pixels lying within specific areas. This is achieved by calibrating image capture with a reference frame. Taking this further, multiple reference settings can be stored to allow for multiple light sources in successive captures, which can then be used to highlight different parts of the object under inspection.

4. Image processing

Look up tables and default correction

Look-up table support provides further mechanisms to deal with issues that arise from lighting and the image sensor. By shifting the gamma of the image via the lookup table, it's possible to make full use of the bit resolution of the image sensor's output stream and optimise the contrast within the image.

Defect correction uses the pixels around a failed one to complete the picture – meaning pixels stuck at one value are not interpreted as issues resulting from lighting conditions on the product being photographed.

Regions of interest

Using techniques such as Wide Dynamic Range creates large file sizes, and as such can only be used in camera systems using high-speed transmission standards. Sony makes this feature available in its 154 fps, 5.1 MP series of Camera Link cameras, **XCL-SG510**. In this, Sony has optimised the sensor specifications, image processing knowhow and the high-speed interface. The XCG-SG510 is actually capable of multiple regions of interest as a result of this combination of technological knowhow.

And while it is possible to deliver wide dynamic range with standards such as GigE, there is a significant data limit imposed, and it is only possible with a lower-end sensor.

One way to counteract this, especially for longer-distance transmission standards like GigE, is by focusing on regions of interest – programming the image sensor to send only portions of an image, and enabling bandwidth to be optimised.

In turn, this allows more imaging subsystems, more sensors and more captures made, so individual processes can be monitored more effectively.

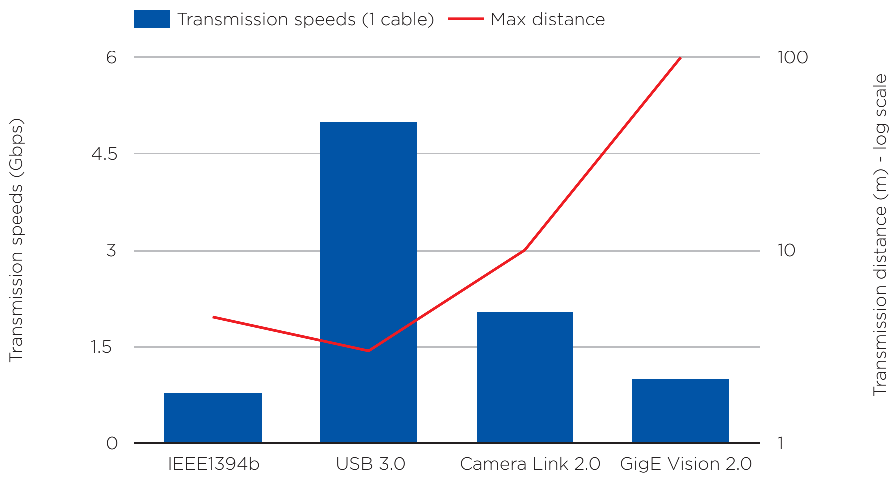
Sony's XCG-CG series of GigE modules are the industry's first cameras able to act as the master in an IEEE1588-controlled system. Their unique acquisition scheduler delivers levels of timing precision that are several orders of magnitude greater than competing devices.

SONY®

5. Data management and transmission

How you stream data is crucial and each transmission standard comes with trade-offs – typically distance vs bandwidth, but also in terms of complexity and cost. In this section we examine the major interfaces and pros/cons of each.

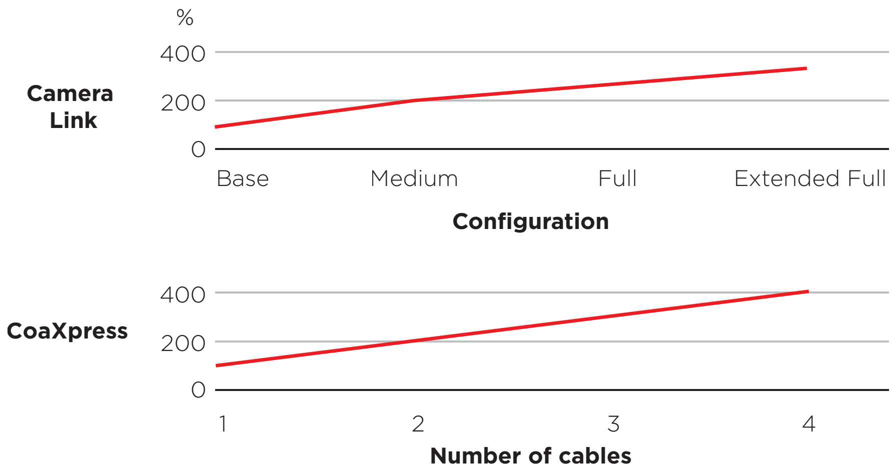
Traditionally, the major standards used in machine vision applications have been GigE, Camera Link and Firewire (IEEE1394b), with USB 3.0 also beginning to be used.



The more common machine vision transmission standards, balancing trade-offs of transmission speed and distance.

5. Data management and transmission

This is, of course, not the complete picture with trade-offs also including cost, the number of cameras that can be placed on a network, the ability to improve transmission speeds by additional cables and GigE's ability to use IEEE1588 to reduce overall system cost (see section 6). For example, under the CoaXPress CXP-6 Quad, 4 cables can be used to deliver 25Gbps (4 * 6.25 Gbps). Camera Link can use two cables, but the total speed depends on the configuration setup and it is not a simple doubling. Furthermore, IEEE1394b enables multiple cables, but for daisy chaining purposes only.



Speed improvements for some standards can be had with multiple cables. With CoaXPress enabling up to 4X on baseline speeds, using 4 cables. Camera Link can use up to two cables, but also allows for further speed gains with configuration changes.

There are however next generation transmission standards which extend both transmission speeds and distance, but the choice of transmission standard also needs to work with existing infrastructure. Camera Link, for example, is commonly used in the Far East; whereas many applications in Europe – especially ITS – have already created systems based on the GigE standard, but are considering Camera Link due to the superior frame rate it supports. However, the implementation of new standards needs to be backwards compatible where possible.

This ability to use existing infrastructure may prove crucial in the success of such standards as 10 GigE. Conversely, Camera Link HS's adoption may take longer as it is not simply an evolution of the 2.0 standard.

SONY®

6. System synchronisation

The IEEE1588 precision timing protocol for GigE v2.0 is enabling systems to be more accurately configured, with accuracies down to the µs.

Be it for agriculture, for ITS functions, for factory robots, indeed for a vast (and growing) number of applications, the ability for multiple cameras – and the entire system – to communicate in order to take more informative images is vital.

For example, a vegetable inspection robot that picks bad fruit from a conveyor (with produce coming down constantly, often covered in dirt) needs images from multiple cameras to interpret exactly which meet supermarket standards, which are acceptable albeit imperfect, which can go to juice, and which are damaged, infected or mouldy.

To manage this, with few false positives and even fewer false negatives, it needs multiple cameras – colour, near-infrared and even polarised and/or hyperspectral – to distinguish between a normal mark, a bruise, an infection and even a hidden object under the skin.

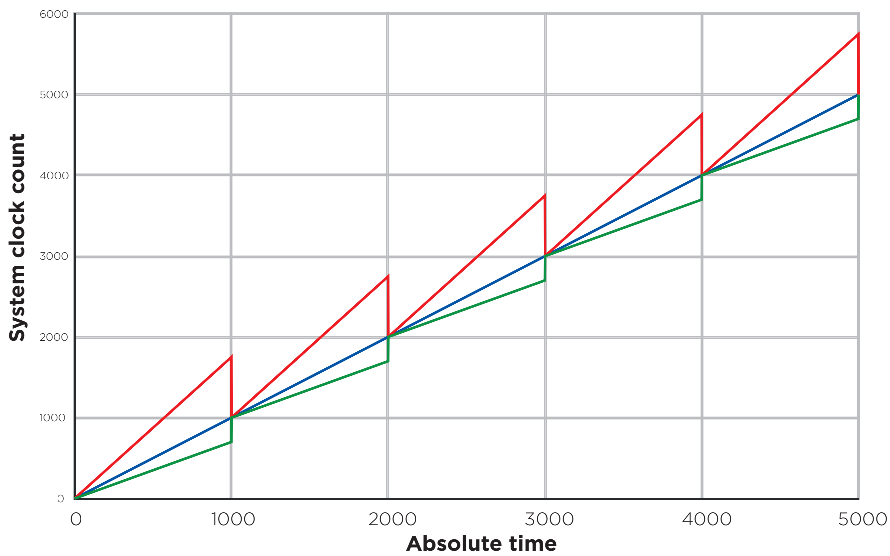
The ability for cameras to fire at precisely the same time and capture exactly the same image is therefore vital.

With a standard camera, there is an internal clock. These clock speeds are arbitrary and unique to each camera and there is no link between devices. As each device on a network has its own clock; you cannot programme a module to fire at a specific time.

Traditionally this has been managed using hardware triggering, based on technologies such as GPS (accurate to the nanosecond), but this adds significant cost and creates a single point of failure in the system.

6. System synchronisation

Synchronisation mechanism



Using the precision timing protocol it's possible to synchronise multiple elements in a system to the master clock and achieve microsecond firing accuracy.

However, the GigE v2.0 communication standard allows the use of the precision timing protocol IEEE1588. This is a software protocol, which dynamically assigns a master clock and, at regular intervals, synchronises all components in the system to the same clock.

In machine vision systems, cameras were traditionally capable of acting solely as the slave device, with a dedicated item of hardware used as master, plus a backup included in case of failure.

Using this protocol it is possible to get the precision down to microsecond accuracy. While less precise than GPS chip's nano-second synchronisation, it is more than accurate enough for virtually all machine vision applications.

Furthermore, the standard is being used to link not just cameras, but the robots themselves, the camera's lens or the lighting system. Individual components can exchange parameters via Ethernet and be triggered over Ethernet, which means it's possible to accurately synchronise, for example, the light pulses and the camera firing and adjust quickly on the fly.

6. System synchronisation

IEEE1588 PTP Function Comparison

Feature	Sony	Competitor A	Competitor B	Competitor C
Free Run Sync. usec precision	**Yes**	**Yes**	**No**	**No**
Master Function Work without external master	**Yes** Preferred time setting	**Yes** Non preferred time setting	**No** External master is necessary	**Yes** Non preferred time setting
Action Command Types Has several hundreds micro-seconds jitter	**Software trigger GPO control Userset load** Can hold 16 commands cue	**Software trigger** Cannot hold commands cue	**Software trigger** Cannot hold commands cue	**Software trigger** Cannot hold commands cue
Acquisition Scheduler Sony original usec precision scheduler	**Yes**	**No**	**No**	**No**

In addition to enabling the camera to become the master, Sony's XCG-CG510 series also uniquely combines an acquisition scheduler as well as software-trigger, GPO-control and UserSet load action command types. This enables SONY's GSCMOS machine vision cameras to realise the most precise pre-scheduled image acquisition synchronisation in real-world environments. Indeed, this level of precision is several hundred times more precise than competing modules.

7. Summary

Recent advances in machine vision technologies have enabled an ever greater ability to deliver images with both exceptional speed and accuracy. By engineering the camera design correctly to take into consideration not just the sensor's capabilities, but the environment it will be working in, it is possible to bring out the key elements of the image and enable better evaluation.

Sony has a prestigious and impressive history in the design, development and production of cameras, extracting the best possible image from its industry-leading sensors. The move to global shutter CMOS has enabled Sony to deliver an even greater array of features to improve accuracy, speed and reduce overall system cost.

Sony's XCL-SG510 Camera Link cameras enables exceptionally-high-speed moving objects to be captured accurately, even in variable light conditions or when shadows and glare create further challenges. The module, designed for use in factory automation and non-manufacturing markets, captures 5.1 MP images at 154 fps as well as including features such as region of interest and wide dynamic range.

Sony's XCG-CG series of GigE cameras are the industry's first cameras capable of not just working as part of a system using the IEEE1588 precision time protocol, but acting as the master in it – eliminating the need for additional hardware, reducing overall system cost. This, combined with their unique acquisition scheduler and and action command type functionality makes them more than two orders of magnitude more precise than competing modules.

The modules also use the leading Sony Pregius IMX264 sensor to deliver 5.1 MP images at 23 fps, with its sister range, the Sony XCG-CG240, delivering 2.4 MP images at up to 41 fps.

Innovations in sensor development and camera design by Sony present ever more possibilities to system integrators, vision system designers and users to improve control, accuracy and recognition in more and more challenging circumstances where camera failure is not an option. It is through this momentum that Sony can continue to offer total cost of ownership advantages to customers.

XCL-SG510

XCG-CG240 & XCG-CG510

For further details on these cameras visit Sony's **image-sensing-solutions.eu** website. Or to find out how Sony's cameras can be used to improve the accuracy of your manufacturing process contact us at **zone@eu.sony.com**