

# A high speed tri-vision system for automotive applications

Marc Anthony Azzopardi · Ivan Grech ·  
Jacques Leconte

Received: 27 May 2009 / Accepted: 2 February 2010 / Published online: 23 February 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

## Abstract

*Purpose* Cameras are excellent ways of non-invasively monitoring the interior and exterior of vehicles. In particular, high speed stereovision and multivision systems are important for transport applications such as driver eye tracking or collision avoidance. This paper addresses the synchronisation problem which arises when multivision camera systems are used to capture the high speed motion common in such applications.

*Methods* An experimental, high-speed tri-vision camera system intended for real-time driver eye-blink and saccade measurement was designed, developed, implemented and tested using prototype, ultra-high dynamic range, automotive-grade image sensors specifically developed by E2V (formerly Atmel) Grenoble SA as part of the European FP6 project – SENSATION (advanced sensor development for attention stress, vigilance and sleep/wakefulness monitoring).

*Results* The developed system can sustain frame rates of 59.8 Hz at the full stereovision resolution of 1280×480 but this can reach 750 Hz when a 10 k pixel Region of Interest (ROI) is used, with a maximum global shutter

speed of 1/48000 s and a shutter efficiency of 99.7%. The data can be reliably transmitted uncompressed over standard copper Camera-Link® cables over 5 metres. The synchronisation error between the left and right stereo images is less than 100 ps and this has been verified both electrically and optically. Synchronisation is automatically established at boot-up and maintained during resolution changes. A third camera in the set can be configured independently. The dynamic range of the 10bit sensors exceeds 123 dB with a spectral sensitivity extending well into the infra-red range.

*Conclusion* The system was subjected to a comprehensive testing protocol, which confirms that the salient requirements for the driver monitoring application are adequately met and in some respects, exceeded. The synchronisation technique presented may also benefit several other automotive stereovision applications including near and far-field obstacle detection and collision avoidance, road condition monitoring and others.

**Keywords** Synchronisation · High-speed automotive multivision · Active safety · Driver monitoring · Sensors

M. A. Azzopardi (✉)

Department of Electronic Systems Engineering,  
Engineering Building, University of Malta,  
Msida, Malta, MSD2080  
e-mail: marc.azzopardi@um.edu.mt

I. Grech

Department of Microelectronics and Nanoelectronics,  
Engineering Building, University of Malta,  
Msida, Malta, MSD2080  
e-mail: ivan.grech@um.edu.mt

J. Leconte

Medical, Industrial & Emerging Imaging BU, E2V,  
Grenoble, France  
e-mail: jacques.leconte@e2v.com

## Definitions, Acronyms, Abbreviations

AIA	automated imaging association
CCD	charge couple device
DLL	delay-locked loop
DSNU	dark signal non-uniformity
FPN	fixed pattern noise
FPGA	field programmable gate array
FVAL	frame valid
HMI	human machine interface
H-Sync	horizontal synchronisation signal
LADAR	laser detection and ranging
LVAL	line valid

LVDS	low voltage differential signal
MTBF	mean time before failure
NIR	near infra red
OCS	occupant classification system
ODS	occupant detection system
OPS	out of position sensing
OWS	occupant weight sensor
PLL	phase-locked loop
PRNU	photo response non uniformity
P-Sync	pixel synchronisation signal
QoS	quality of service
RADAR	radio detection and ranging
ROHS	reduction on hazardous substances
ROI	region of interest
SODAR	sonic detection and ranging
TWI	two wire interface
V-Sync	vertical synchronisation signal

## 1 Introduction

Over the coming years, one of the areas of greatest research and development potential will be that of automotive sensor systems and telematics [1, 2]. In particular, there is a steeply growing interest in the utilisation of multiple cameras within vehicles to augment vehicle Human-Machine Interfacing (HMI) for safety, comfort and security [3].

For external monitoring applications, cameras are emerging as viable alternatives to systems such as Radio, Sound and Light/Laser Detection and Ranging (RADAR, SODAR, LADAR/LIDAR). The latter are typically rather costly and either have poor lateral resolution or require mechanical moving parts [4].

For vehicle cabin applications, cameras outshine other techniques with their ability to collect large amounts of information in a highly unobtrusive way. Moreover, cameras can be used to satisfy several applications at once by re-processing the same vision data in multiple ways, thereby reducing the total number of sensors required to achieve equivalent functionality. However, automotive vision still faces several open challenges in terms of optoelectronic-performance, size, reliability, power consumption, sensitivity, multi-camera synchronisation, interfacing and cost.

In this paper, several of these problems are addressed. As an example, driver head localisation, point of gaze detection and eye blink rate measurement is considered for which the design of a dash-board-mountable automotive stereovision camera system is presented. This was developed as part of a large FP6 Integrated Project - SENSATION (*Advanced Sensor Development for Attention, Stress, Vigilance and Sleep/Wakefulness Monitoring*). The overarching goal of

SENSATION was to develop non-invasive sensors, including stereovision cameras, for general human vigilance monitoring. Stereovision methods offer unique advantages for automotive applications and in this case they permit the extraction of many cues that allow driver vigilance to be reliably quantified.

The system presented here employs a novel method of addressing the synchronisation problem that arises in such system. It also demonstrates a novel method for reliably transporting high speed, synchronised, stereovideo over a single Camera-Link® interface. By virtue of its simplicity, this method is also presented as a means to reduce the overall cost of high performance stereovision systems. The ability of multiplexing stereovideo onto a single Camera-Link® cable halves the cabling cost as well as the impact on a vehicle's cable harness weight. This method is readily extendable to multivision systems [5–8].

The camera system is built around a matched set of prototype, ultra-high dynamic range, automotive-grade, image sensors specifically developed and fabricated by E2V Grenoble SA for this application. The sensor which is a novelty in its own right, is the AT76C410ABA CMOS monochrome automotive image sensor. This sensor implements a global shutter to allow distortion-free capture of fast motion. It also incorporates an on-chip *Multi-ROI* feature with up to eight Regions Of Interest (ROI) with pre-programming facility and allows fast switching from one image to another. In this way, several real-time parallel imaging processing tasks can be carried out with one sensor. Each ROI is independently programmable *on-the-fly* with respect to integration time, gain, sub-sampling/binning, position, width and height.

A fairly comprehensive series of “bench tests” were conducted in order to test the validity of the new concepts and to initially verify the reliability of the system across various typical automotive operating conditions. Additional rigorous testing would of course be needed to guarantee a mean time before failure (MTBF) and to demonstrate the efficacy of the proposed design techniques over statistically significant production quantities.

## 2 Application background

The set of conceivable automotive camera applications is an ever-growing list with some market research reports claiming over 10 cameras will be required per vehicle [9]. The incomplete list includes occupant detection, occupant classification, driver recognition, driver vigilance and drowsiness monitoring [10], road surface condition monitoring, intersection assistance [11], lane-departure warning [12], blind spot warning, surround view, collision warning, mitigation or avoidance, headlamp control, accident record-

ing, vehicle security, parking assistance, traffic sign detection [13], adaptive cruise control and night/synthetic vision (Fig. 1).

## 2.1 Cost considerations

The automotive sector is a very cost-sensitive one and the monetary cost per subsystem remains an outstanding issue which could very well be the biggest hurdle in the way of full deployment of automotive vision. The supply-chain industry has been actively addressing the cost dilemma by introducing Field Programmable Gate Array (FPGA) vision processing and by moving towards inexpensive image sensors based on Complementary Metal Oxide Semiconductor (CMOS) technology [14]. Much has been borrowed from other very large embedded vision markets which are also highly cost-sensitive: These are mobile telephony and portable computing. However, automotive vision pushes the bar substantially higher in terms of performance requirements. The much wider dynamic range, higher speed, global shuttering, and excellent infra-red sensitivity are just a few of the characteristics that set most automotive vision applications apart. This added complexity increases cost. However, as the production volume picks up, unit cost is expected to drop quite dramatically by leveraging on the excellent economies of scale afforded by the CMOS manufacturing process.

Some groups have been actively developing and promoting ways of reducing the number of cameras required per vehicle. Some of these methods try to combine disparate applications to re-use the same cameras. Other techniques (and products) have emerged that trade-off some accuracy and reliability to enable the use of monocular vision in scenarios which traditionally required two or more cameras [10, 15, 16]. Distance estimation for 3D obstacle localisation is one such example. Such tactics will serve well to contain cost in the interim. However, it is expected that the cost of the imaging devices will eventually drop to a level where it will no longer be the determining factor in

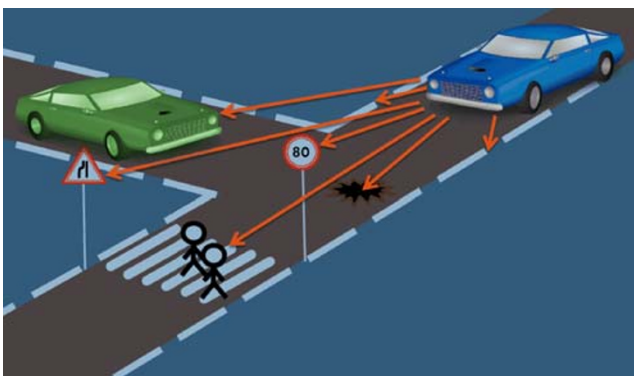


Fig. 1 Some automotive vision applications

the overall cost of automotive vision systems. At this point, we argue that reliability, performance and accuracy considerations will again reach the forefront.

In this paper the cost issue is addressed, but in a different way. Rather than discarding stereo- and multi-vision altogether, a low-cost (but still high-performance) technique for synchronously combining multiple cameras is presented. Cabling requirements are likewise shared, resulting in a reduction in the corresponding cost and cable harness weight savings.

## 2.2 The role of high speed vision

A number of automotive vision applications require high frame-rate video capture. External applications involving high relative motion such as traffic sign, oncoming traffic or obstacle detection are obvious candidates. The need for high speed vision is perhaps less obvious in the interior of a vehicle. However, some driver monitoring applications can get quite demanding in this respect. Eye-blink and saccade measurement, for instance, is one of the techniques that may be employed to measure a driver's state of vigilance and to detect the onset of sleep [10, 16]. It so happens that these are also some of the fastest of all human motion and accurate rate of change measurements may require frame rates running up to several hundred hertz. Other applications such as occupant detection and classification can be accommodated with much lower frame rates but then the *same* cameras may occasionally be required to capture high speed motion for visual-servoing such as when modulating airbag release or seatbelt tensioning during a crash situation.

## 2.3 A continued case for stereovision/multivision

Several of the applications mentioned, stand to benefit from the use of stereovision or multivision sets of cameras operating in tandem. This may be necessary to extend the field of view or to increase diversity and ruggedness and also to allow accurate stereoscopic depth estimation [11]. Then, of course, multivision is indeed one of the most effective ways of counteracting optical occlusions.

Monocular methods have established a clear role (alongside stereoscopy) but they rely on assumptions that may not always be true or consistently valid. Assumptions such as uniform parallel road marking, continuity of road texture, and operational vehicle head or tail lights are somewhat utopian and real world variability serves to diminish reliability. Often, what is easily achievable with stereoscopy can prove to be substantially complex with monocular approaches [17]. The converse may also be true, because stereovision depends on the ability to unambiguously find corresponding features in multiple views.

Stereovision additionally brings a few challenges of its own, such as the need for a large baseline camera separation, sensitivity to relative camera positioning and sensitivity to inter-camera synchronisation.

Not surprisingly, it has indeed been shown that better performance (than any single method) can be obtained by combining the strengths of both techniques [18, 19]. As the cost issue fades away, monovision and multivision should therefore be viewed as complimentary rather than competing techniques. This is nothing but yet another example of how vision data can be processed and interpreted in multiple ways to improve reliability and obtain additional information.

In this paper, the benefit of combining stereo and monocular methods is demonstrated at the hardware level. A tri-vision camera is presented that utilises a synchronised stereovision pair of cameras for 3D head localisation and orientation measurement. Using this information, a third monocular high-speed camera can then be accurately controlled to rapidly track both eyes of the driver using the multi-ROI feature. Such a system greatly economises on bandwidth by limiting the high speed capture to very small and specific regions of interest. This compares favourably to the alternative method of running a stereovision system at high frame rate and at full resolution.

#### 2.4 The importance for high synchronisation

One of the basic tenets of multivision systems is the accurate temporal correspondence between frames captured by the different cameras in the set. Even a slight frequency or phase difference between the image sampling processes of the cameras would lead to difficulties during transmission and post processing. Proper operation usually rests on the ability to achieve synchronised, low latency video capture between cameras in the same multivision set. Moreover, this requirement extends to the video transport mechanism which must also ensure synchronous delivery to the central processing hubs. The need for synchronisation depends on the speed of the motion to be captured rather than the actual frame rate employed, but in general, applications which require high speed vision will often also require high synchronisation.

Interestingly, even preliminary road testing of automotive vision systems reveals another sticky problem – *camera vibration*. This is a problem that has already been faced many years ago by the first optical systems to enter mainstream vehicle use [20] – The optical tracking mechanisms used in car-entertainment CDROM/DVD drives are severely affected by automotive vibration and fairly complex (and fairly expensive) schemes are required to mitigate these effects [21].

The inevitable vibration essentially converts nearly all mobile application scenarios into high speed vision problems

because even low amplitude camera motion translates into significant image motion. The problem gets worse as the subject distance and/or optical focal length increases.

Mounting the cameras more rigidly helps by reducing the vibration amplitude, but it also automatically increases the vibration frequency which negates some of the gain. Active cancellation of vibration is no new topic [22]; however, this usually comes at a disproportionate cost. Thus, while high frame rates may not be important in all situations, short aperture times and high synchronisation remain critically important to circumvent the vibration problem.

A small numerical example quickly puts the problem into perspective. Consider a forward looking camera for in-lane obstacle monitoring based on a  $\frac{1}{4}$  inch,  $1024 \times 512$  image sensor array with an active area of  $5.7 \times 2.9$  mm behind a 28 mm (focal length) lens. If such a system is subjected to a modest 10 mrad amplitude, sinusoidal, angular vibration at 100 Hz, simple geometric optics implies a peak pixel shift rate of around 32,000 pixels/sec.

Thus, if the error in correspondence between left and right stereo frames is to be limited to a vertical shift comparable to one pixel, a stereovision system would require a frame synchronisation accuracy which is better than 30 microseconds. Then on the road, the levels of vibration can get significantly worse and this does not yet take into account the additional high speed motion that may be present in the field of view. In summary, synchronisation is a problem that has been largely overlooked and will become more important as the industry and consumer performance expectations increase.

In this paper, a synchronisation technique based on matched cameras sharing a single clock is presented. The system affords a very high degree of synchronisation – in fact, much higher than is actually demanded by the driver monitoring application. Synchronisation difficulties arising during initialisation and camera mode changes are also addressed in this paper using a novel frozen-clock programming technique.

#### 2.5 High bandwidth interconnect and processing

Automotive vision faces another formidable challenge – bandwidth. Having several cameras running at high frame rates and at high resolutions quickly pushes such applications into the multi GBit/s domain. This poses new pressures on a sector that is still barely warming up to multi-MBit/s interface speeds. New automotive video interface standards will be required, and while it makes sense to base these on existing and proven interconnects, it may be argued that a completely new standard is needed to properly address the requirements of this peculiar market. The stage is set for a standards-war and in fact, one is



currently brewing which should eventually see the evolution of a veritable *Automotive Video Bus*. Such a bus faces a tall order which includes: low cable cost, low interface cost, low specific weight, multi-GBit/s sustained throughput, multiplex-ability, preservation of synchronisation, high integrity, excellent electromagnetic compatibility (EMC) characteristics, low latency, low jitter, and a minimum 5 m cable span without repeaters [23].

There is of course a second repercussion of such high bandwidths. Impressive data rates necessitate equally impressive computational power in order to perform all the associated video processing *in real-time*. This is fairly problematic considering the limited capabilities of most automotive embedded processors, but this is changing with the entry of FPGAs into the automotive market [23–25]. Aside from offering substantial (and sufficient) in-line processing power, FPGAs also serve to reduce cost by combining most of the interface glue-logic into a single chip. Then, FPGAs have the added appeal of re-configurability which allows aftermarket updates through simple firmware changes – though this raises several security concerns [25].

### 3 Video interfaces

A survey of currently available interface standards reveals that none of the present offerings are ideally suited to faithfully transport high speed, high resolution, synchronised stereovideo over appreciable distances. The following is a comparative discussion of the merits and shortcomings of the various interfaces.

#### 3.1 Bandwidth considerations

The Interface throughput is the major concern since high resolutions are desirable and the required frame rates can reach into the high hundreds per second. At a moderate 200 frames per second, a 10 bit per pixel, greyscale,  $640 \times 480 \times 2$ , stereovision system generates video at 1.229 GBit/s. Even  $1536 \times 768 \times 2$  at 12 bit is not at all farfetched for certain applications and this touches 5.662 GBit/s which is impossible to accommodate on most current interfaces. Evidently, the interface is a bottleneck that needs to be addressed.

For our driver monitoring application, 60 Hz is sufficient for accurate head localisation. However 200 Hz or more is desirable for fast eye-saccade and eye-blink capture. Running the entire system at 200 Hz at full resolution is therefore wasteful. By using a trinocular system, the frame rate of the stereovision pair can be set to 60 Hz, while a third monocular camera tracks the eyes alone at 200 Hz using a pair of 10,000 pixel ROIs. This way, assuming

10bit, the bandwidth requirements are reduced to a more manageable (369+40) MBit/s. The information collected using the stereovision system guides the ROI placement for the third camera.

Hence, for this application, the strict requirement is for an interface that can sustain 409 MBit/s of throughput. However, in view of the possibility of other vision applications and future resolution improvements, the design should aim for an interface which should be able to handle a significantly higher bandwidth.

#### 3.2 Latency and jitter considerations

Throughput alone does not fully describe the problem. Low system latency is another aspect that cannot be neglected. Practically all of the automotive vision applications mentioned, depend on real-time low latency access to the processed output from the vision information. The driver vigilance application is no exception but other even more demanding applications come to mind. At 90 km/h a vehicle covers 25 m every second. A single second of lag in a high speed obstacle detection situation can make the difference between avoiding an accident and reacting too late. The problem with latency is that it all adds up. There is latency at the sensor, transmission latency, processing latency and actuator (or human) latency. If this totals up to anything more than a few tens (or hundreds) of milliseconds, the effectiveness of most of these safety systems would be seriously compromised. Of course, establishing an exact value for the desired latency is no precise science because it depends on the situation.

Video processing is perhaps the most important contributor to the overall latency and this usually needs dedicated hardware to keep up with the demands. FPGAs were already mentioned in this respect. Transmission is next in line in terms of latency severity. Delays due to buffering should be minimised or eliminated. Moreover, the latency should be fixed and uniform. Many signal processing techniques and control systems do not react too well to random variations in their sampling interval. Hence, there is a strong requirement for deterministic system behaviour with negligible transmission and processing time jitter.

#### 3.3 Video interface selection

Analogue interfaces were once the only practical way of transmitting video information. The analogue bandwidth of coaxial copper cables is fairly good, latency is minimal and such interfaces offer excellent temporal determinism. Multi-camera support is also readily possible using radio frequency (RF) modulation/multiplexing and is a mature and reliable technique. However, guaranteeing signal

integrity of analogue video becomes prohibitively difficult at high resolutions and frame rates. Moreover, with the prevalent use of intrinsically digital CMOS image sensors, it would be highly inconvenient and expensive to convert digital video data to analogue and back just for transmission. The future lies entirely with digital. Table 1 provides a comparative summary of the various interfaces that were considered in this project.

The initial obvious choice for digital video transmission technology is to look at established standards in the consumer electronics market. This could exploit the associated economies of scale and high maturity. However, a closer look reveals several shortcomings. While serial packet-transport protocols such as the Ethernet-derived GigE-Vision standard can sustain up to 750 Mbit/s [26], they have poor temporal characteristics, including high latency, poor determinism and substantial timing jitter making them rather unsuitable for high performance vision applications [27]. Even so, such throughput is only possible by using Jumbo Framing (a non-standard proprietary

technology) [28]. Central processor (CPU) utilisation can also be unacceptably high.

Multimedia-oriented protocols such as the Universal Serial Bus (USB2) and Firewire (IEEE1394b) only partially address these problems through the inclusion of special isochronous modes of operation. The raw bandwidth is fairly high at 480 MBit/s and 3.2 Gbit/s respectively. However, their timing accuracy is limited to no better than  $\pm 125 \mu\text{s}$ , [29, 30]. Moreover, synchronous transport of multimedia streams over intrinsically asynchronous protocols poses complexities that outweigh the benefits [31].

On the other hand, parallel video bus standards such as RS-422 and RS-644 which are based on parallel Low-Voltage Differential Signalling (LVDS), exhibit low latency, are highly deterministic, are synchronous and are relatively jitter-free by design. They also offer good throughput. Of course, the downside of any parallel bus is a severe limitation in length due to cable delay skew as well as the need for thick expensive cables.

**Table 1** Comparison of some interface standards

Interface Type	Cable Type	Bandwidth (MBit/s)	Temporal Determinism	Multi-Camera	Cable Cost (&Weight)	Complexity
Analogue interfaces						
RGB	CP	GHz	5	N	H	L
Composite	CC	GHz	5	N	M	L
RF Composite	CC	GHz	5	Y	M	L
Component	CC	GHz	5	N	M	L
Automotive media interfaces						
MOST	FB	23	4	Y	L	M
APIX	STP	1000	5	N	M	M
IDB-1394	STP	400	2	Y	L	H
FlexRay	UTP	20	4	Y	M	H
Field busses						
D2B	CT, FB	5.6	3	–	M, L	M
LIN	UTP	0.02	2	–	L	M
CAN2.0	UTP	1.0	3	–	L	H
PROFIBUS	UTP	12.0	3	–	L	M
SPI, I <sup>2</sup> C	UTP	0.1–1.0	2	–	L	M
Consumer-oriented serial interfaces						
Gig-Ethernet	UTP, FB	1000	1	Y	L, L	H
IEEE1394b	STP	3200	3	Y	M	H
USB2	STP	480	3	Y	M	H
Industrial vision parallel interfaces						
RS644	CP	<1000	5	N	H	L
RS422	CP	<1000	5	N	H	L
Industrial vision hybrid interfaces						
Camera-Link	CT, FB	7140	5	N	M, L	L

1 = Poor, 2 = Fair, 3 = Medium, 4 = Good, 5 = Excellent, L = Low, M = Med, H = High, Y = Yes, N = No

CP Parallel Copper, CC Copper Coax, UTP Unshielded TP, STP Shielded TP, FB Fibre

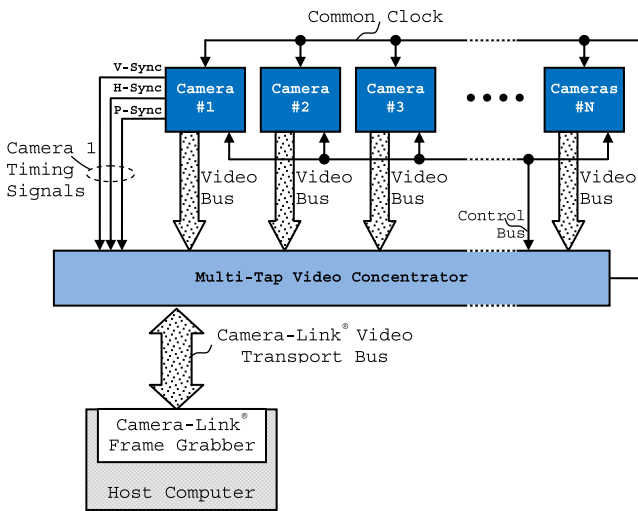


Fig. 2 General multivision system architecture [5]

The automotive industry has a fairly long history of data bus use and development and standards abound, such as the CAN-Bus (CAN2.0), LIN, SPI, D2B, I<sup>2</sup>C and other field busses. The problem common to all of these standards is that they are mostly intended for control applications and real-time, low data-rate sensor interrogation. While determinism is fairly good, the total bandwidth is too low. So while it may be theoretically possible to hook multiple cameras to such busses, in reality, the addition of a single high performance camera would swamp out all the bus resources and it would still not suffice.

The automotive industry and its supply chain have reacted to this clear need for faster and more capable interfaces and there are several new initiatives appearing on the market. FlexRay is a fairly new bus designed to replace the CAN-bus and enable new functionality such as drive-by-wire, high-performance power-trains, safety systems, active suspensions or adaptive cruise control. The Media Oriented Systems Transport (MOST) is primarily designed for consumer multimedia-interconnect such as navigation equipment and in-vehicle entertainment. It claims to be reasonably deterministic from ground up. However, for both these interfaces, the 20 Mbits/s of bandwidth is a non-starter for high speed vision applications. Several compa-

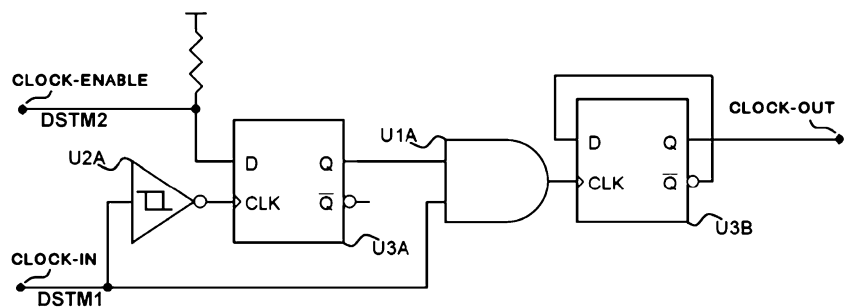
nies have pushed for the adoption of IDB-1394 which is an automotive variant of the highly successful consumer-product: IEEE1394. However this suffers from most of the same problems of its forerunner.

Inova Semiconductor, has made substantial headway with its “Automotive Pixel Link” (APIX®) technology [32, 33] which follows on its GigaStar consumer-oriented interface. This is an asymmetric point-to-point data transport system that is based on serialiser/deserialiser technology and as such promises high throughput, low latency and excellent determinism. As such it straddles the parallel/serial interface domains and offers some of the advantages of both. This is an interesting technology and if the costs can be contained it could gain popularity in the vision market.

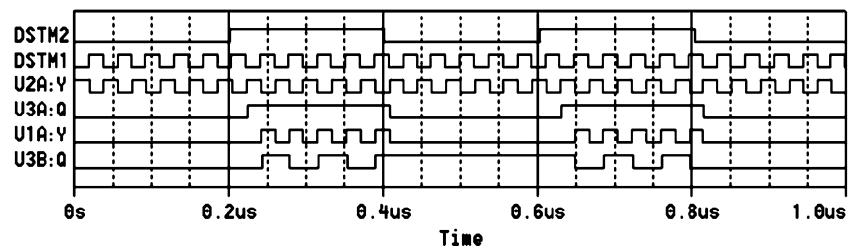
Then finally there is Camera-Link®, which is a proven dedicated machine-vision interface developed by some of the major players in the machine vision market [34]. This also straddles the parallel/serial domains and derives the best benefits from each; having the performance, simplicity and Quality of Service (QoS) of a parallel bus while keeping the desirable cabling benefits of a serial bus. Fibre optic implementations of Camera-Link® take the length limit to the kilometre range [35, 36] and of course fibre implementations offer galvanic isolation, heat/fire resistance and the lowest possible specific weight. Camera-Link® is essentially a unidirectional point-to-point protocol with minimal control bandwidth dedicated to the reverse path but this suits machine vision applications well.

Camera-Link® and APIX® share many technical characteristics that make them ideal for automotive vision although they are intended for different domains. However, they both seem to lack an obvious way for interconnecting multiple cameras per interface. This is where this paper makes a contribution. In this project Camera-Link® was selected as a basis for what could become an Automotive Video Bus due to the reasons mentioned in the forgoing as well as its superior bandwidth. Camera-Link was extended, to allow the interconnection of multiple synchronised cameras in a multivision set. APIX® would have been an equally adequate starting point but APIX® compliant hardware is only just appearing on the market. That said,

Fig. 3 A clock gating circuit



**Fig. 4** Glitchless operation of clock gating circuit



much of what is presented for Camera-Link® is also directly applicable to APIX® so the results are portable across both interfaces.

#### 4 Overview of synchronisation techniques

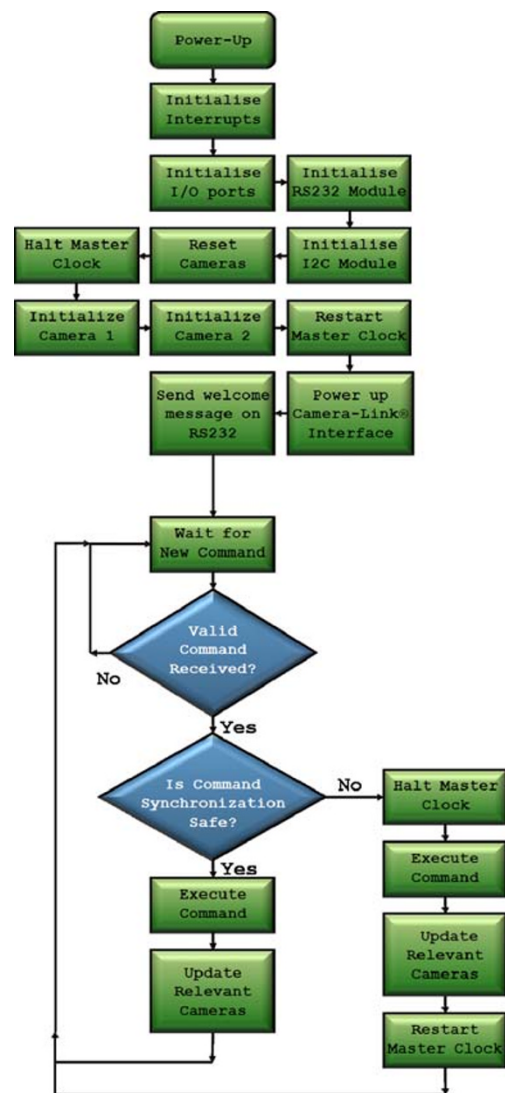
As already mentioned, the effective application of stereo-vision or multivision systems depends on the ability to capture synchronised video from two (or more) separate locations. There is of course the possibility of using beam splitting optics and a single camera [37], but this can be exceedingly cumbersome and expensive and as such precludes applications needing substantial viewpoint separation. On the other hand, solving the problem using multiple cameras to generate and transmit synchronised video signals is non-trivial and there have been numerous attempts to address it, as evidenced by the several related patents.

The oldest methods of synchronisation between multiple cameras date back to the 1980's when the 'genlock' (generator lock) principle [38], became commonplace for use in video broadcasting houses, video editing and special effects [39]. This was, and still is, quite adequate for TV broadcast systems. However, as the frame rates and pixel rates increase, it fails due to the transportation lag incurred in transferring a genlock signal between cameras. Electro-mechanical synchronisation techniques were also proposed [40], but quickly fell into disfavour as electronics gradually took over all aspects of this field.

Some techniques rely on post processing (frame shifting) to achieve synchronisation. The relative frame lag is measured either by comparing recorded motion present in the two video streams [41, 42], or by actively inserting artificial optical cues into the field of vision of the cameras [43]. This avoids the need for explicit synchronisation and is touted as a means of reducing costs but there are a number of scenarios where the net complexity and cost is increased by the need of the additional post-processing step. Moreover, this technique is not universally applicable such as in cases where there is no motion in the captured sequences or where interference with the scene is not acceptable. This method of synchronisation is additionally severely limited in the accuracy it can achieve since the

resulting video sequences could still be misaligned by as much as half the inter-frame duration, on average.

Schemes that involve the transfer of vertical or horizontal or synchronisation pulses between the cameras in a multivision system, [44, 45], have similar shortcomings to the Genlock concept, from which they are derived. Phase-Locked Loops (PLLs) and Delay-Locked Loops (DLLs) can be used to compensate for delays but this adds significant complexity and ultimately limits the pixel clock



**Fig. 5** Command marshalling by a camera controller



rate. Store and forward techniques proposed by the same authors [46] allow synchronous transmission of video data, but do nothing to guarantee synchronous frame capture. They also add complexity and the cost of a large high-speed buffer, and unavoidably introduce a small but distinct latency in the delivery of the video data which may be a significant disadvantage for certain high speed applications.

### 5 System architecture

The stereovision system implemented and presented here was meant to demonstrate the feasibility of achieving a steady stream of high speed, precisely synchronised stereo-video over a standard interface when using typical off-the-shelf CMOS automotive-grade image sensors (represented by the AT76C410).

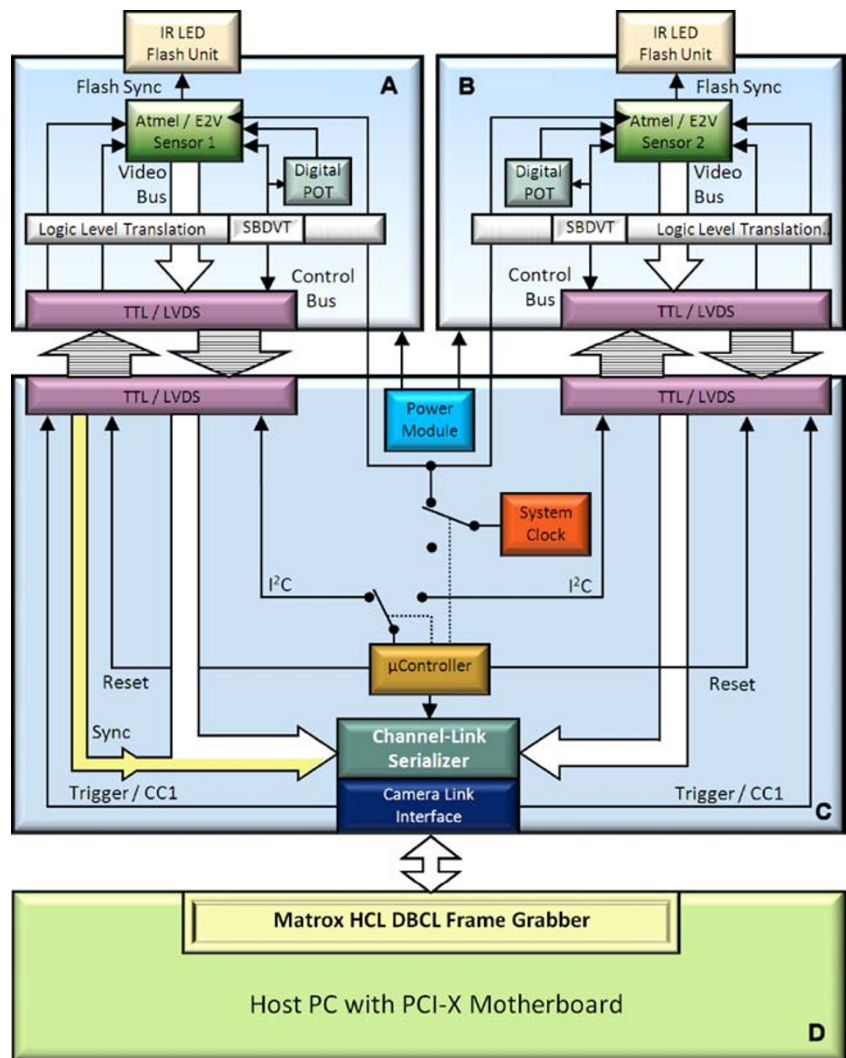
The proposed method involves the use of *matched* cameras or image sensors, which are driven by a common

clock as well as operate under identical operating conditions thereby guaranteeing an identical internal state and synchronised output timing behaviour. Compared with other synchronisation techniques, this significantly reduces latency and again keeps the costs to a minimum while lending itself for a complete solution.

Flexibility, minimal weight, low latency, high performance, high reliability and low overall cost were the major objectives of this undertaking.

To this effect, the generic architecture shown in Fig. 2 is proposed. Any number of identical cameras can be symmetrically connected to a central video concentrator. The cameras are perfect replicas of one another (matched to within close tolerances in terms of the electronics) and the image sensors are taken from matched sets that have been produced in the same fabrication run (from the same silicon wafer) to guarantee equivalent performance and timing characteristics when supplied with a common clock. To further reduce variability even the cables connecting the

**Fig. 6** A stereovision implementation [7]



cameras to the concentrator board are of matched length and composition. Hence, matching is largely a design consideration and should not significantly impact the production cost of such systems. Accurate electrical matching is important to ensure the temporal alignment of all timing signals.

The video concentrator has a number of roles, the most important being that of ensuring that every camera is operating under the same programmatic and electrical conditions at all times and its internal architecture conforms to this principle at every level. Another role is that of combining as many video streams as possible at an early stage before transmission across the vehicle to a central processor. This reduces the quantity (and weight) of cabling.

## 6 Clock modulation

A major challenge often encountered in such situations is the need to simultaneously initialise or re-program all the cameras in the system. This is quite problematic considering that the majority of CMOS image sensors are configured over relatively slow serial interfaces (often on shared bus). In practice commands have to be sequentially delivered to each of the cameras and for certain commands this process would invariably result in frame/line phase misalignment between the cameras.

This problem has been neatly resolved by recognising that most CMOS image sensors are fully static state machines. This allows their clock to be halted and restarted at will, without any lasting consequences on the state. In addition, these CMOS sensors do not require the master clock to be active in order to access and reprogram the internal control registers. For programming, a separate clock, which has no effect on the sensor state, can be delivered via their I<sup>2</sup>C interface. Thus, before delivering commands to the image sensors, the common master clock can be halted. This conserves the machine state. Only after all the commands are sequentially sent to all the cameras, is the clock re-started. The overall effect is equivalent to having reconfigured all the cameras at the same instant.

However, not all camera commands require such a procedure. Some commands do not affect synchronisation at all and it may even be desirable, in certain cases, to be able to apply arbitrary operating parameters to different cameras without interrupting the video capture. One such example is a change in pixel gain and/or integration time. Thus, the solution adopted in this design involves marshalling all the commands and distinguishing between those that are synchronisation safe from those that are not. Only those commands that affect synchronisation are intercepted for halted-clock execution.

A camera controller residing in the video concentrator module controls the delivery of the common master clock to the cameras by means of a clock gating circuit. This clock gating circuit is capable of synchronously interrupting and reconnecting the clock without causing any glitches at the output that might adversely affect the sensor state.

The clock gating circuit, shown in the schematic of Fig. 3, takes a clock and a clock-enable line as inputs. This input clock must run at twice the frequency required by the cameras. When the clock-enable line is held at logic low, the AND gate U1A isolates the output D-flip-flop U3B

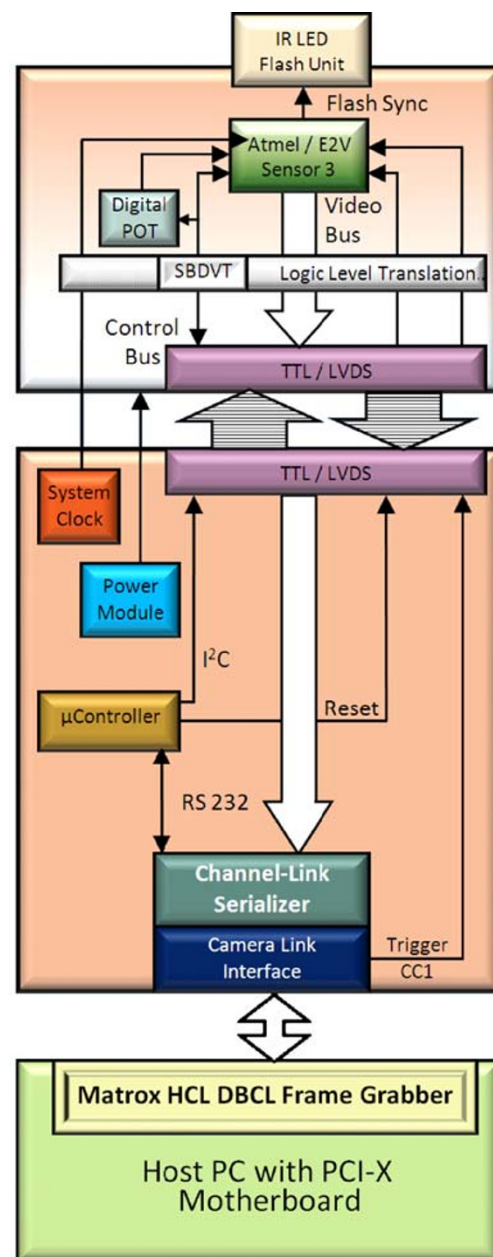
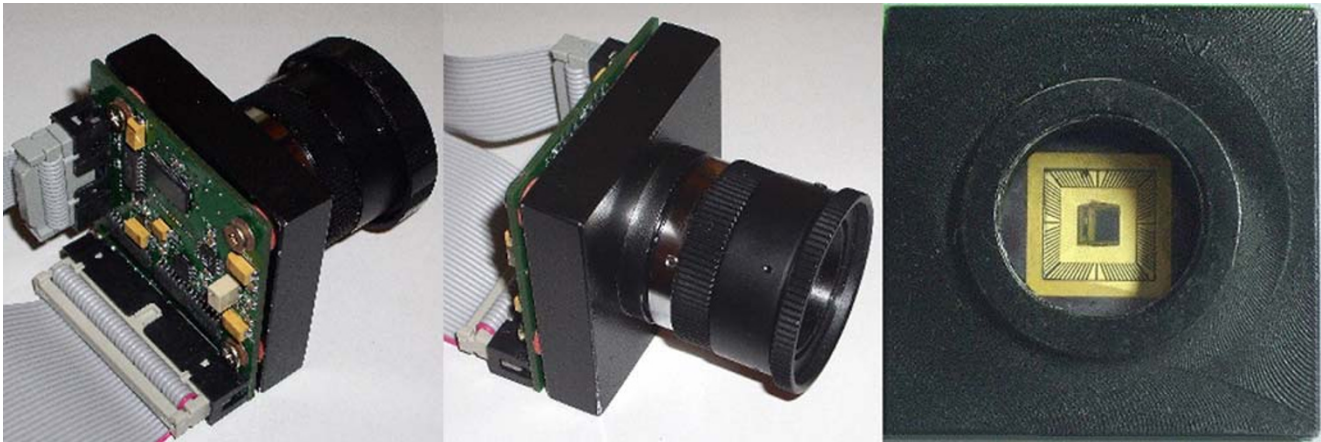


Fig. 7 Independent monovision camera [7]



**Fig. 8** The camera modules

which holds its last held state, interrupting clock transfer. When the clock-enable line is held high, the AND gate U1A relays the clock to the output D-flip-flop U3B which divides the frequency, producing a clean 50% duty cycle clock. The negative edge triggered D-flip-flop U3A only conducts changes in the clock-enable line to the AND gate U1A at the negative edges of the incoming clock which satisfies set-up time requirements of the output flip-flop U3B.

Referring now to the simulation result shown in Fig. 4, several signals are shown describing the operation (as a function of time) of the clock gating circuit when supplied with clock signal DSTM1:1 and clock-enable line signal DSTM2:1. U2B:Y shows the inverted clock which is fed into D-flip-flop U3B. U3A:Q shows the re-synchronised clock-enable line pulse. U1A:Y shows the gated clock. U3B:Q shows the gated output of the circuit after frequency division.

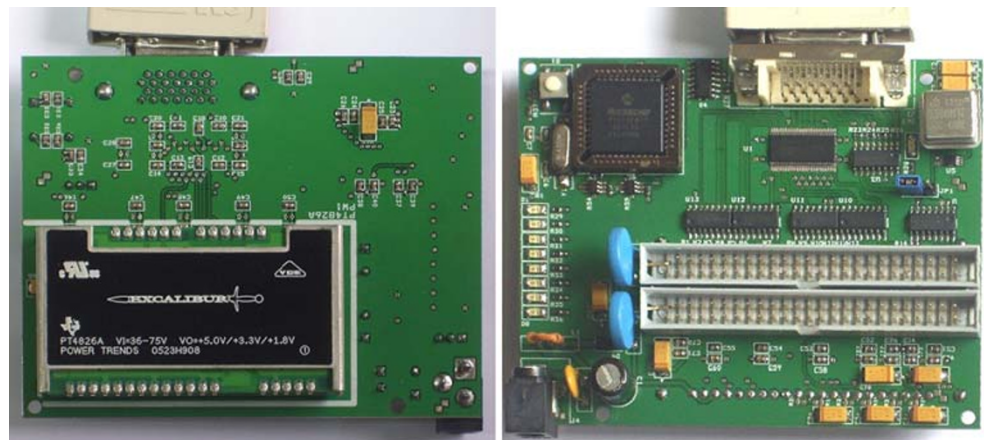
The camera controller consists of a low cost 8-bit Microchip PIC16F877A microcontroller embedded into the video concentrator. The selection of micro-controller is immaterial so long as it possesses the required RS232 and

I<sup>2</sup>C interfaces. It is programmed to execute the flowchart shown in Fig. 5, which is here described in terms of the stereovision implementation of the proposed system, but is easily extended to systems involving more than two cameras. This flowchart represents a simple but novel method for preserving synchronised camera behaviour during the power up sequence and also during any configuration changes performed in the cameras.

After power-up, the controller initialises the interrupt handler and enables or disables the relevant interrupts in the microcontroller. Next, the I/O ports are initialised followed by the initialisation of the RS232 and I<sup>2</sup>C hardware ports. Next, the cameras are reset by issuing a reset pulse on the dedicated camera reset lines. At this point, the clock is halted in preparation for the initialisation of the two cameras. The initialisation of the second camera is performed after the initialisation of the first camera, but this does not pose a problem so long as the clock remains halted. Then the clock is restarted and the Camera-Link<sup>®</sup> interface is powered-up.

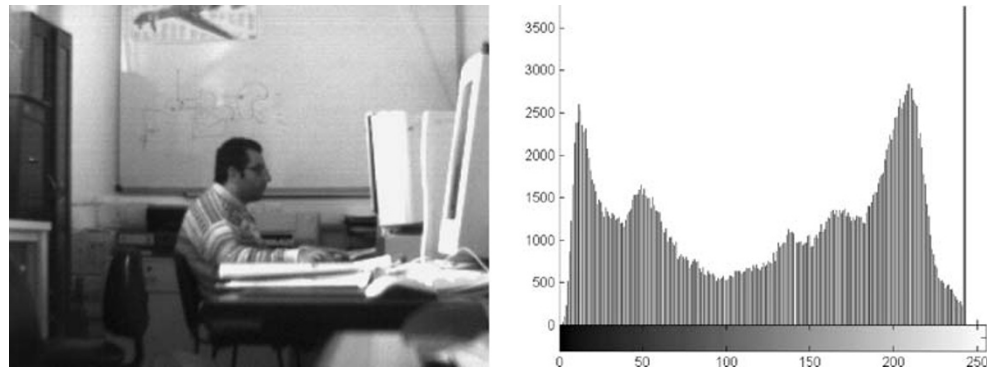
After sending a welcome message over RS232, the controller enters into a wait state. If a command is received

**Fig. 9** The stereovision video concentrator module





**Fig. 10** Histogram test results for normal operation



during this time, it is first validated and if it is not found to be valid, the controller discards it and re-enters the wait state. If the command is on the other hand, valid, the command is accepted and classified depending on whether it is synchronisation safe or not. If it is synchronisation safe, it is executed and the cameras are updated.

If the command is not synchronisation safe, the clock is halted, the command is executed, the relevant registers within both cameras are updated and finally the clock is restarted. After completion of command processing, the camera controller re-enters the wait state in order to accept new commands.

## 7 Video multiplexing

The second major role of the video concentrator module is to multiplex the video streams onto a single interface. It starts by collecting the video data from each camera in the stereovision pair, which at this point can be assumed to be in near perfect synchronism. The corollary of this is that the frame, line and pixel synchronisation signals from all the cameras are practically indistinguishable and all but one can effectively be discarded.

In order to multiplex the video streams over a single interface, the video concentrator emulates a multi-tap video source to simultaneously transmit all the streams together with a single set of synchronisation signals. This exploits

the fact that most off-the-shelf machine vision frame grabber hardware is already equipped to handle and demultiplex multi-tap video [47]. The classic way of transporting multi-tap video was to have parallel data links. However, this defeats the light-weight and low-cost objectives. A different method is therefore required.

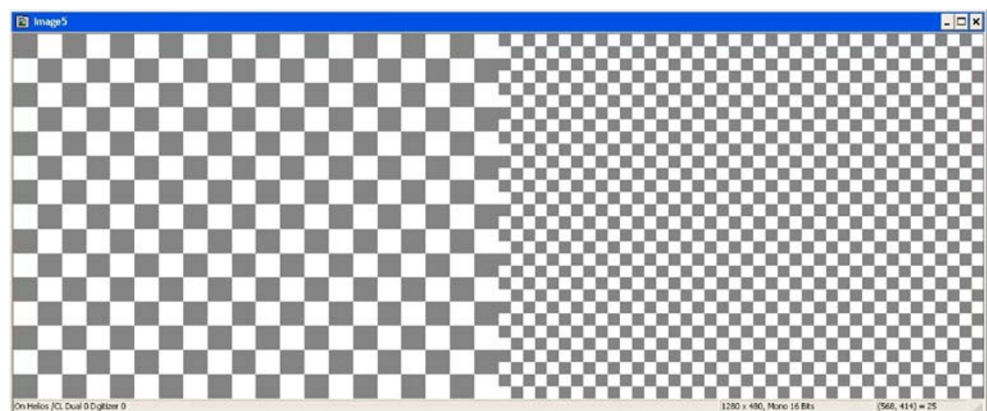
Camera-Link<sup>®</sup> natively caters for multi-tapping and the official specification already defines several modalities for transporting multi-tap video over a single interface. Provided that the video streams are in perfect synchronism, as would be the case had they come from a real multi-tap camera, they can be transmitted over Camera-Link<sup>®</sup> without any additional processing or buffering. In the case of APIX<sup>®</sup>, the three primary colour (RGB) channels of a “virtual colour” camera can be used instead of multi-tapping to the same end.

The drawing in Fig. 6 shows, some architectural detail of the stereovision camera system. It comprises two cameras (A and B), a stereovision video concentrator (C), a Camera-Link<sup>®</sup> cable, a Camera-Link<sup>®</sup> frame grabber, and a host computer (D).

As previously mentioned the cameras are identical in every respect. The left camera is operated as a master while the right camera is operated as a slave but this distinction is merely the result of the way the outputs from the cameras are treated by the video concentrator.

Each camera comprises a CMOS image sensor that triggers an LED flash unit using a dedicated flash sync

**Fig. 11** Multi-tap video multiplexing test





**Fig. 12** Indirect synchronisation test results



pulse. The image sensor generates Transistor-Transistor-Logic (TTL) timing signals and drives a video bus while it accepts a clock, an I<sup>2</sup>C serial control bus and a TTL camera reset signal. The cameras are connected to the video concentrator with a high integrity bidirectional LVDS link which carries the video bus and the timing signals towards the concentrator and carries the camera reset and control bus towards the cameras. TTL to LVDS transceivers at both ends, perform the conversion in both directions.

The video concentrator comprises, amongst other things, a common master clock, a clock gating circuit, a camera controller, a Channel-Link<sup>®</sup> serialiser and a Camera-Link<sup>®</sup> Interface. The Channel-Link<sup>®</sup> serialiser takes the two video busses and the Camera-Link<sup>®</sup> timing signals and serialises them onto four high speed differential serial lines. These are then mapped onto the Camera-Link<sup>®</sup> interface (in the order defined by the standard) and finally transmitted over the Camera-Link<sup>®</sup> cable to the frame grabber. The host computer ultimately receives and de-multiplexes the video data to produce a wide 1280x480 composite stereo-image.

One of the requirements of the driver monitoring application was the ability to observe the driver's eyes closely at very high frame rate. This was needed in order to be able to extract the driver's blinking rate and saccade movements with sufficient temporal resolution. For this, the

ROI feature was employed which allows small regions of a few thousand pixels to be sampled at several hundred hertz. A third separate camera (Fig. 7) was needed to allow it to be decoupled from the stereovision pair.

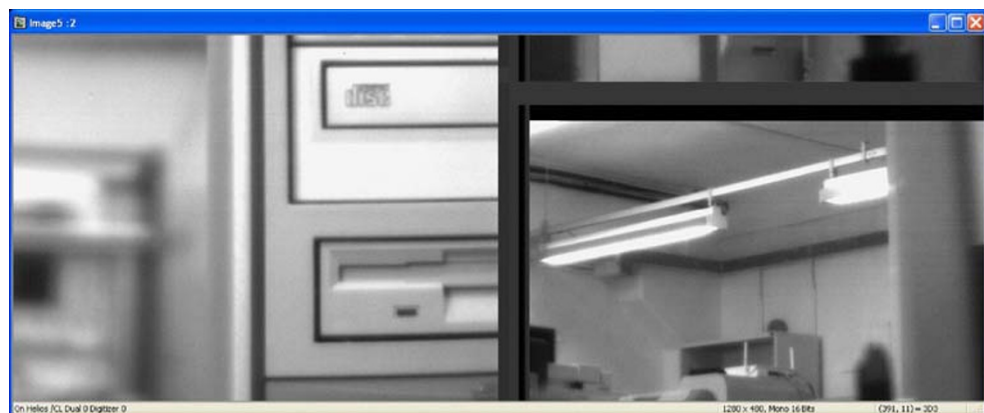
This third camera was connected to the same frame grabber via the secondary Camera-Link base channel, which also provides a completely independent control path.

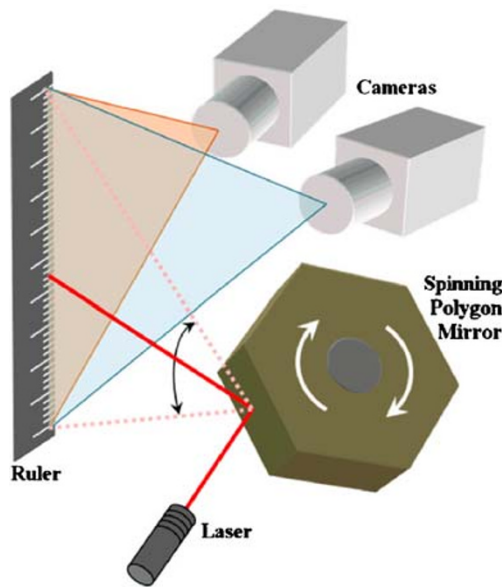
## 8 Implementation results

The stereovision system was implemented using the following core components:

- E2V (formerly Atmel) AT76C410AB Prototype Automotive Image sensors
- Arizona Microchip PIC1LF877A 8-Bit flash micro-controllers
- National Semiconductor DS90LV048ATM LVDS to TTL Receivers
- National Semiconductor DS90LV047ATM TTL to LVDS Transmitters
- National Semiconductor DS90CR287MTD 28-Bit 85 MHz ChannelLink<sup>®</sup> Serialisers
- Texas Instruments Excalibur PT4826N DC/DC Converters

**Fig. 13** Experimental control showing picture tear





**Fig. 14** Laser polygon scanner experiment

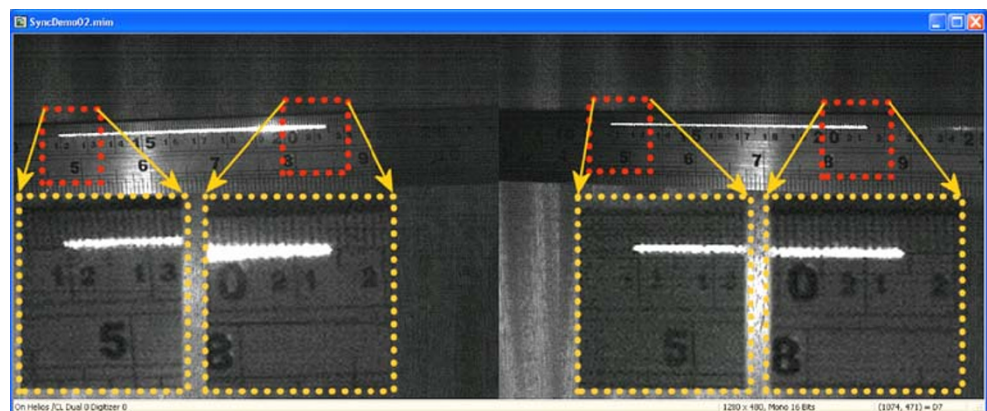
All system modules were assembled in-house on 6 layer PCBs that were fabricated at Beta Layout GmbH. The camera controller was programmed in a hybrid C/ASM language. Figure 8 shows photographs of the finished camera modules while Fig. 9 shows the video concentrator.

## 9 Testing philosophy

The design process was completed over three iterations and four complete prototype copies of the final design were produced and delivered to other partners in the project. Although these were prototypes, some measure of quality had to be assured. Testing was carried out over five stages to comprehensively assess different aspects of the tri-vision camera system.

The first tests focused on the quality of the design, board-fabrication and assembly processes. These ensured that the final systems were free from manufacturing

**Fig. 15** Laser scanner test results

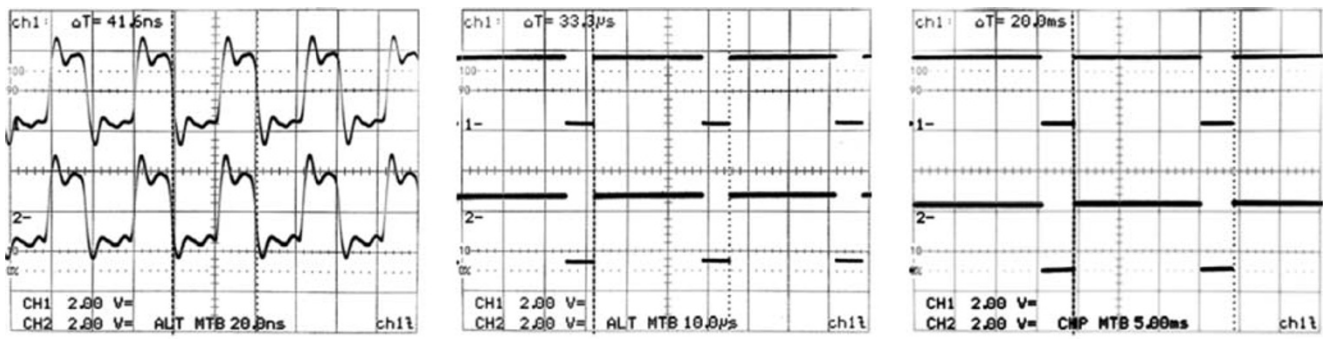


defects. Defects were identified and corrected. The second set of tests focused on the primary objective of the project - that of achieving unconditional precision synchronisation and efficient video multiplexing. These tests validated the novel concepts developed during this project. A third level of tests established the firmware's stability. All the software residing in the camera controller was meticulously tested and every possible execution path was verified to be able to guarantee stability in most scenarios.

The image sensors were prototypes themselves and included numerous novel features and performance attributes applicable to the automotive scenario. These had to be specifically tested and verified against the manufacturer's expected behaviour [48]. E2V Grenoble SA conducted an extensive series of in-house tests to establish the validity of their product against a set of pre-agreed acceptability criteria. A selection of these tests was again repeated at a system level. Finally the optical performance of the cameras was assessed and the data collected was used to perform fine adjustments to obtain focus uniformity and optical axis alignment. The level of testing was necessarily limited to bench tests due to the statistically insignificant number of cameras produced. The primary objective behind the testing was to validate the design concept and to weed out potential manufacturing defects. Higher production volumes would permit more rigorous forms of testing.

## 10 Testing methods and results

Histogram tests are one of the most effective diagnostic methods for camera circuits. These quickly provide insight into the integrity of the entire video data path. Any stuck bits are quickly manifested as periodic gaps in the histogram. The periodicity of the gaps indicates the affected bit while the orientation (right or left handed) indicates the type of fault (stuck at high or stuck at low respectively). For



(a) P-Sync (20ns/div) (b) H-Sync (10µs/div) (c) V-Sync (5ms/div)

Fig. 16 Direct electrical synchronisation results

an  $X$  bit image, the periodicity  $P$  of the histogram artefact indicates the affected bit  $B$  where:  $B = X - \log_2(P)$ .

Figure 10 shows the normal histogram of a complex image captured with one of the cameras.

Video multiplexing tests were initially demonstrated without the use of any cameras. A chequer-board test image generator was constructed using a system of counters on a Field Programmable Gate Array (FPGA) and the ensuing data was fed into a Channel-Link<sup>®</sup> serialiser, emulating a multi-tap video source.

This in turn, delivered the test video streams to a frame grabber. The resulting images were carefully analysed for picture tears and jitter but none were detected. Figure 11 is a screen shot of the received test stereovision image as demultiplexed by the frame grabber.

Synchronisation tests were performed directly and indirectly. The latter method of testing consisted in simply operating the stereovision system while connected to a frame grabber. Such a setup is fairly sensitive to synchronisation and is a quick way of ensuring compliance. If the phase difference between the two cameras exceeds half a pixel period, it would cause easily detectable picture tears. Figure 12 shows a stereovision capture test result, and as can be observed, no such picture tears are present.

This should then be compared with a control test in which the clock gating function was deliberately disabled during the initialisation sequence. Figure 13 shows the expected resulting picture tear in the slave camera image (left half).

A rather more scientific method for directly demonstrating accurate synchronisation consisted in the simultaneous capture of a fast moving object against a reference background. However, for adequate sensitivity, the object had to move at km/s rates and the only practical method found for achieving this was by reflecting an intense (100 mW) collimated laser beam off a rapidly spinning polygon mirror onto a ruled surface. The polygon mirror spins on a synchronous drive which means that the angular velocity may be accurately determined. With this method, a precise  $7.736 \text{ kms}^{-1}$  scanning velocity was achieved which

on a 1.0 mm ruled surface gave a temporal resolution of 130 ns. The experimental setup is depicted in Fig. 14.

The result achieved is shown in Fig. 15. This image, shows the laser scan line sweeping past a steel ruler as captured by the left and right cameras. Enlarged inlays (in red borders) showing the salient parts of the scan line (in yellow borders) are shown below the ruler as indicated by the arrows. As expected, the locations of the start and end points of the laser scan-line in the left and right stereo images matches perfectly.

The temporal resolution of optical methods is limited. In this case, the reason is that these images were taken at the shortest aperture time of the cameras ( $1/48000 \text{ s}$ ) and if the scan velocity is increased any further, it becomes impossible to fit a complete scan line within the camera’s field of view, which in turn makes it impossible to simultaneously compare the duration, (start and end time) of each aperture interval.

However, having established that the cameras are optically synchronised, better resolution can be obtained with electrical methods. An oscilloscope can be used to directly compare the video synchronisation pulses generated by the two cameras in the pair. The slightest synchronisation misalignment would immediately be apparent as a phase difference between these pulses. Figure 16 shows the

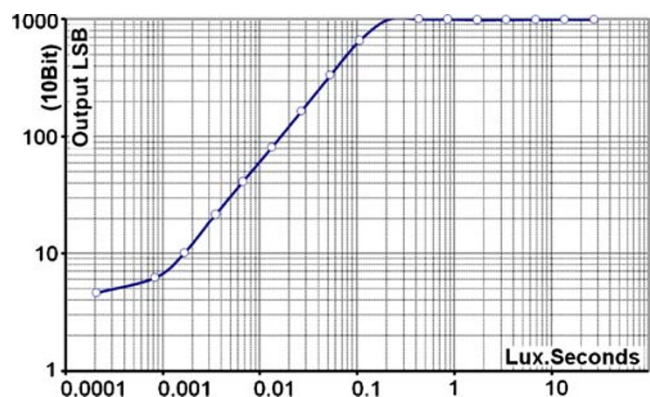


Fig. 17 Nominal photo-response test results



**Fig. 18** Before and after non-linearity compensation



oscilloscope test results for the pixel (a), horizontal (b) and vertical (c) synchronisation signals respectively. The top traces pertain to the master camera while the bottom traces are derived from the slave. The phase difference between the traces was again beyond measurement using a 2.5 GS/s oscilloscope with a 10 fold rate of oversampling and stood at much less than 100 ps.

Image sensor performance was tested in a number of ways. The sensors were engineering samples and the tests were mostly intended to check whether these prototypes were operating as expected, and also to ensure that the overall camera design is well behaved in all conditions.

A test which is particularly relevant to the automotive scenario is the operation of the system at extreme temperatures and with non-ideal configurations such as unequal cable lengths and non nominal supply voltages. The system was successfully operated at temperatures ranging from  $-20^{\circ}\text{C}$  to  $+120^{\circ}\text{C}$  in non condensing environments. Such tests are by no means accurate or conclusive, but they do offer an added level of confidence in the quality of the prototypes. In a production environment such products would of course subjected to lengthy thermal and power cycling to establish long-term reliability. However, this was beyond the scope of the project.

The Nominal Photo-response Characteristic of the cameras was measured directly using a Mastech LX1330B Digital Luxmeter. A 75 W tungsten-filament incandescent lamp at a colour temperature of 2820 K was used as a

reference light source. The luminous exposure (in Lux. seconds) was modulated by adjusting the distance between the source and the cameras, by using mesh filters and finally by altering the total integration time at the sensors. This gave a wide enough range for luminous exposure. Figure 17 shows the resulting response.

The photo-response characteristic was linear for the most part but non-linear at the higher light levels. This combination permitted excellent behaviour at normal illumination levels but at the same time it extended the dynamic range to allow the cameras to handle direct sunlight. This is a distinguishing feature between automotive-grade image sensors and other sensors. Figure 18 shows the resulting images before (left) and after (right) compensation for the nonlinear characteristic.

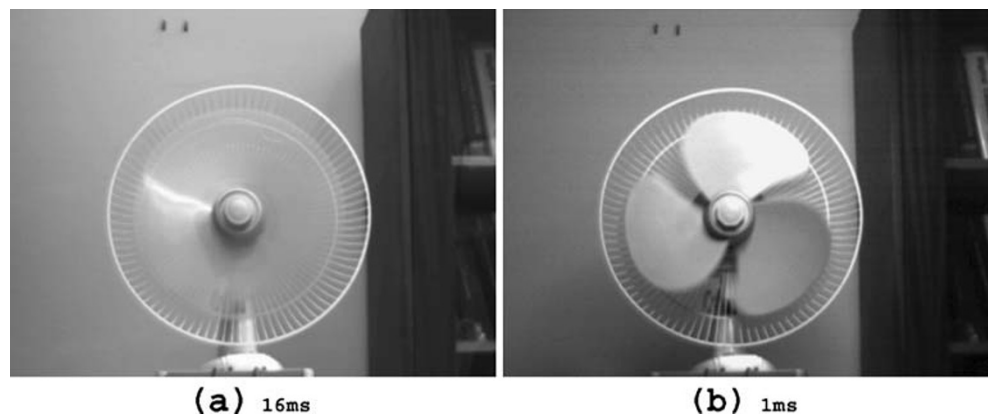
The image sensors feature an adjustable dynamic range. This gives them the capability to alter the partitioning between the linear and nonlinear portion of their photo-response characteristic by externally controlling the pixel bias voltage and allows the user to sacrifice linearity in return for better dynamic range performance. This trade-off parameter can also be rapidly adjusted in real-time, thus allowing machine vision algorithms to optimise the dynamic range depending on the operating circumstances. With this technique the dynamic range was effectively extended to a remarkable 123 dB [49] which compares favourably to previous reports [30].

The advantage of an adjustable dynamic range is clearly demonstrated in a particularly challenging scenario as

**Fig. 19** Dynamic range test results

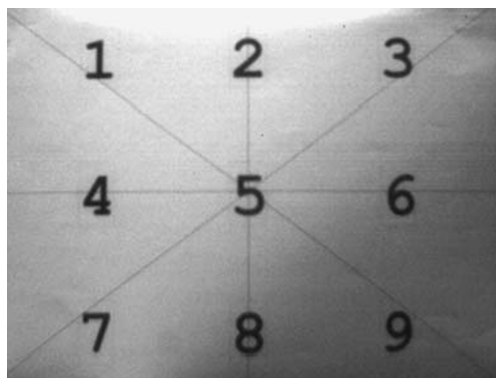




**Fig. 20** High speed operation at full resolution

shown in Fig. 19, where a modestly illuminated background is contrasted with a bright fluorescent lamp shining directly into the camera lens. Both images are taken using identical exposure conditions (integration time and gain). However, the left image was taken with the camera running with its nominal dynamic range showing severe over-exposure. On the other hand, the image on the right is obtained after a dynamic range adjustment. The result shows clearly distinct background and foreground features with little, if any, over-exposure.

High speed operation is mandatory for capturing fast eye and eyelid movements. Motion blur and motion distortion are not acceptable in this application. This automatically requires very short aperture times and the use of a global shutter. A sustained frame rate of at least 200 Hz and an integration time as short as 1 ms were important design criteria. These features were tested using a *fan test* in which a rapidly spinning fan propeller was imaged under various conditions. Figure 20 shows such a fan spinning at its maximum speed of 1311 RPM imaged once with an integration time of 16 ms and another time with an integration time of 1 ms. At this rotation rate, the peripheral velocity of the fan blades is 19.9 m/s. The cameras support integration times as short as 20.8  $\mu$ s but a 1 ms aperture should result in measurable motion blur spanning just under 2 cm. This matches what is observed in practice. No motion distortion is observed.

**Fig. 21** ROI test target image

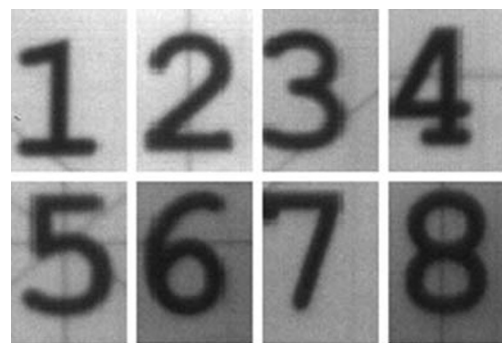
In order to allow very high frame rates without overwhelming the internal image sensor Analogue to Digital Converter (ADC) with samples and the host computer with data, a special (ROI) mode is included. This restricts the field of view to a small portion containing the object of interest and can be resized and shifted in real-time to track the object. The reduced number of pixels allows substantially higher frame rates to be achieved - up to 750 Hz for 10,000 pixels.

As mentioned previously, the ROI feature is particularly useful for the third monocular camera that is being used for high frame rate tracking of the driver's eyes. However it can also be used in the stereovision pair provided that the aspect ratio and size of the ROI is set identically in both cameras.

The cameras also allow sequential tracking of multiple ROIs – up to 8 ROIs can be defined. Figure 21 shows a test target image and Fig. 22 shows its decomposition in 8 consecutive frames of a  $50 \times 70$  pixel ROI, using the Multi-ROI feature. The camera cycles indefinitely over all the active ROI frames, potentially feeding up to eight separate image processing routines in tandem.

## 11 Summary of results

Various other results and system characteristics are summarised in Table 2.

**Fig. 22** 8-way multi-ROI decomposition

The stereovision system was finally deployed for driver vigilance monitoring in a luxury test vehicle, the “Lancia Thesis 2.4 20 V Emblema”, at FIAT, Turin and was then tested successfully at the Centre for Research and Technology Hellas (CERTH) in Thessaloniki. Figure 23 shows a photo of some of the equipment installed in this vehicle.

## 12 Discussion

This paper addresses the synchronisation problem which arises in high speed multivision camera applications. In this paper, a novel precision synchronisation method is pre-

sented which exploits the similarity of behaviour and performance of matched cameras (or image sensors) by subjecting them to a common clock. By managing their operating conditions, it can guarantee an identical internal state and synchronised output timing behaviour, which will in turn permit the combined transmission over great lengths over a single high performance vision interface.

Reports of comparable systems are fairly scarce and poorly documented with respect to the synchronisation problem. A system based on the Fillfactory (now Cypress Semiconductor) FUGA-1000 random-access image sensor was developed by the Graz University of Technology in Austria [30]. This system appears to allow concurrent

**Table 2** Summary of results

Parameter	Result achieved
Resolution	640×480 Progressive scan
Output format	10 bits digital
Sensor fabrication technology	0.18 μm CMOS monochrome
Resolution	640×480 Progressive scan
Optical format	1/3"
Colour depth	10 bits monochrome
Pixel size	6 μm×6 μm
Pixel rate	Max 27 MHz
Integration time	20.8 μs up to 1.36 s
Optical dynamic range (non-lin)	123 dB
Sensor power supply (Anlg/Dig)	3.3 V / 1.8 V
Spectral range	350–1,050 nm
Electronic shutter	Global Shutter
Anti-blooming feature	Yes
Region of interest (ROI) mode	Yes
Multiple ROI mode	Yes: 8-Way
Sensor configuration interface	I <sup>2</sup> C
Camera configuration	Software
Camera configuration interface	Serial: RS232
Camera frame rate (full format)	59.8 fps Max @ 24 MHz
Camera frame rate (10 k ROI)	750 fps Max @ 24 MHz
Camera pixel rate	24 MHz (max 27 MHz)
Image transport lag	1 frame duration
Configuration interface speed	9,600 or 19,200 Baud
Camera dimensions (W × H × D)	54×54×37 mm <sup>3</sup>
Video interface	Single Base Camera-Link™
Safely aspects	Over-voltage, over-current, polarity-inversion protected
Stream synchronisation	< 1 ns (<< 1 pixel clock cycle)
Power supply	36 V to 75 V dc
Power consumption (at 50 fps)	3.44 W
Image sensor package	CLCC 84
Lens port	C-Mount
Lens focal length	12 mm
Lens aperture	f/1.3
Operating temperature	0° to +40°C

image capture from two cameras. However, no detail is reported on any synchronisation technique or its temporal performance, nor is any reference made to any method of synchronous sensor programming or initialisation and how this can be managed during resolution or frame-rate changes. These aspects are thoroughly studied and adequately addressed in our paper.

Our method not only addresses the issue of generating accurately synchronised video signals in a simple and very economical way, but also avoids the need for transferring frame or line synchronising pulses between cameras. This avoids the delays associated with the transmission of such pulses making it applicable to systems requiring ultra-high speed operation without posing any serious restrictions on the relative positioning of the cameras. Much higher frames rates can be realistically achieved this way. In addition, the high precision synchronisation afforded by our method allows the aggregation of multivision cameras into a system that mimics a multi-tap camera. This allows the combined and faithful transmission of the outputs of several cameras over a single Camera-Link<sup>®</sup> connection over substantial distances. The method presented here extends, without violating, the provisions for multi-tap video, as laid out in the Camera-Link<sup>®</sup> specification. This method also avoids the need for a “store and forward” mechanism and hence does not incur any of the cost, complexity and latency associated with the internal buffering used in other methods.

The selection of Camera-Link<sup>®</sup> offers important advantages for the high speed transfer of highly synchronised stereovideo. Indeed, the Graz University system, which was based on the popular USB2 serial interface, faced significant temporal non-uniformity and bandwidth limitations, as described by Muehlmann et al. in [30]. USB2 presented a bottleneck and hampered the full exploitation of the image sensor’s capabilities. This is due to the inflexible 125  $\mu$ s microframe USB2 time base. Moreover, the need to transmit



**Fig. 23** System installed in a Lancia Thesis Emblema.

10-bit pixel data over the 16-bit wide peripheral interface of USB2 also put the designers in a quandary, by having to choose between truncating the 2 least significant bits of each sample or having to face a 6 over 16 bit bandwidth wastage. On the other hand, the 32-bit Camera-Link<sup>®</sup> bus width allows up to 3 10-bit pixels words (from synchronised cameras) to be accommodated with minimal bandwidth wastage.

The Graz University system uses an FPGA for glue logic integration. However, it also needs a fairly large FPGA to accommodate all the image sensor addressing logic, the USB interface logic and to eternally manage its stereovision ROI function. In contrast, our system places a very flexible Multi-ROI function and all of the associated pixel addressing onto the image sensor, which greatly simplifies the external glue logic required. An FPGA is therefore not essential although the use of a small FPGA or ASIC (Application-Specific Integrated Circuit) would result in reduced size and power consumption.

The system being presented offers an excellent dynamic range of 123 dB which compares well with other contemporary image sensors such as the Fillfactory FUGA-1000 [30]. However, the addition of an adjustable dynamic range offers the unique ability to match the sensor’s sensitivity to the image being captured in real-time.

### 13 Further work

The demonstration system developed is of course an experimental prototype in many respects and future work can place all of the interface logic into an FPGA or ASIC which will reduce size and power consumption by a further 80% at the very least. In addition, during the course of the development of this tri-vision imaging system, E2V has developed an improved imaging sensor EV76C560, based on the AT76C410. These sensors offer a step change in performance and versatility and pave the way for much improved automotive cameras and new applications.

This new device has enhanced ROI features including individual header, footer and inbuilt image histogram computation. The latter facilitates the fast computation of auto-exposure. Though the number of regions of interest has been reduced to four, this was found to be adequate for most automotive applications. Each ROI can be read from 1 to 256 times. The ROIs can now operate in two modes: Multi-Integration, Multi-Readout (MIMR) and Single-Integration, Multi-Readout (SIMR). With SIMR, all four ROIs are captured during the same integration interval and are thus guaranteed to be synchronized. On the other hand, MIMR allows each ROI to be sampled sequentially, which guarantees uniform sampling latency.

The resolution has now been increased to 1280 $\times$ 1024 pixels with a pixel rate of 114 Mpixels/s. The sensor has

5 T pixels and can operate in either global shutter mode or electronic rolling shutter (ERS) mode. Higher gain has been implemented in the pixel output amplifiers, resulting in higher low light sensitivity.

Excellent dynamic range can be obtained with a new bi-frame integration technique. This offers the flexibility of separately integrating dark and bright regions of wide dynamic range. Such a feature is especially useful when combined with real-time High Dynamic Range Imaging (HDRI) [50–52] and compositing techniques such as Blendfest® [53] to produce exceptionally wide dynamic ranges.

The new sensor can be configured at high speed via the use of a Serial Peripheral Interface (SPI) bus, which is accessible even during standby mode. Thus real-time and frozen-clock configuration remains possible.

## 14 Conclusions

The system presented here offers a complete, high accuracy and high performance video multiplexing solution for multivision applications in general. The system was designed, built and tested for the automotive environment and was also built around the latest automotive image sensors, making it as realistic to the application as practically possible.

Higher resolutions, high frame rates and high accuracy are of critical importance for automotive vision [2, 11]. New developments by sensor manufacturers and the rising number of demanding applications sitting on the horizon (awaiting better cameras) indicates that the market will be performance-driven for the foreseeable future. Such performance needs to be reflected at the systems-level and hence, the objective of this paper was to definitively address the synchronisation problem that arises between different cameras when combined in a multivision set.

However, another significant contribution is the very substantial reduction in the cabling required to connect multivision cameras to central hubs through the use of synchronous multiplexing. In this paper, Camera Link® was chosen as the video transport protocol, but the technique is equally applicable to newly emerging high performance video interfaces such as APIX® [33]. The end result is a significant saving in terms of weight and cost. This method makes it possible to break new barriers in this regard which will again be particularly attractive in the automotive sector.

**Acknowledgments** This project was partially funded by the EU through the IST-507231 SENSATION project. I wish to acknowledge the SENSATION project consortium for their valuable contributions to this work.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Turner JD, Austin L (Feb. 2000) Sensors for automotive telematics. *Meas Sci Technol* 11(2):R58–R79, Berkshire, UK
2. Hock U (23 Jul. 2009) CCD/CMOS Cameras: eyes for cars. New Business Development for Sharp Microelectronics Europe, Design article, Automotive DesignLine Europe, Munich, Germany
3. Beecham M (Jul 2008) Global market review of driver assistance systems - forecasts to 2014. 2008 edition, Just Auto, Chapter 3, Technical review. Aroq Ltd., Worcestershire, UK
4. Hoffmann I (Jan. 2006) Replacing radar by an optical sensor in automotive applications. *Advanced Microsystems for Automotive Applications 2005*, Springer Berlin Heidelberg, Book Chapter, Berlin, Germany, pp 159–167
5. Azzopardi MA. Method for synchronising stereovision or multi-vision camera systems for combined transmission over Camera-Link®. Malta Patent, No: MT#4230, University of Malta, Filed Sep. 2008, Valletta, Malta
6. Azzopardi MA. Method and Apparatus for Generating and Transmitting Synchronized Video Data, PCT Patent Application, No: PCT/EP2009/061553, University of Malta, Filed Sep. 2009, Geneva, Switzerland
7. Azzopardi MA. Stereovision system design using Camera-Link® for Low Voltage Automotive CMOS Image Sensors, M.Phil. Thesis, University of Malta, Submitted Sep. 2008, Msida, Malta
8. Azzopardi MA (Nov. 2008) Camera-Link® and synchronism in automotive multi-vision systems. *Conference Proceedings, 4th International Conference on Automotive Technologies, ICAT-2008, Istanbul, Turkey, Vol. 1*, pp 344–353
9. ABI Research (2007) Camera-based automotive systems regional forecasts and key competitive assessment for driver assistance technology. *Market Research Report, RR-CBAS*, ABI Research, New York, USA
10. Smith P, Shah M, da’Vitoria Lobo N (Sep. 2000) Monitoring head/eye motion for driver alertness with one camera. *Conference Proceedings, 15th International Conference on Pattern Recognition, ICPR-2000, Barcelona, Spain, Vol. 4*, pp 636–642
11. Nedeveschi S et al (Apr. 2009) Stereovision-based sensor for intersection assistance. *Advanced Microsystems for Automotive Applications 2009*, (Smart Systems for Safety, Sustainability, and Comfort), Springer Berlin Heidelberg, Book Chapter, Berlin, Germany, pp 129–163
12. Schubert R et al (Apr. 2009) Lane recognition using a high resolution camera system. *Advanced Microsystems for Automotive Applications 2009*, (Smart Systems for Safety, Sustainability, and Comfort), Springer Berlin Heidelberg, Book Chapter, Berlin, Germany, pp 209–227
13. Luth N et al (Apr. 2009) Lane departure warning and real-time recognition of traffic signs. *Advanced Microsystems for Automotive Applications 2009*, (Smart Systems for Safety, Sustainability, and Comfort), Springer Berlin Heidelberg, Book Chapter, Berlin, Germany, pp 267–285
14. Fossum ER (1997) CMOS image sensors: electronic camera-on-a-chip. *IEEE Trans Electron Devices* 44(10):1689–1698
15. Saxena A, Chung SH, Ng AY (2006) Learning depth from single monocular images. *Stanford University. Adv Neural Inf Process Syst (NIPS)*, (18):1161–1168
16. Bretzner L, Krantz M, Goteborg S (Oct. 2005) Towards low-cost systems for measuring visual cues of driver fatigue and inattention



- in automotive applications. Smart Eye AB, Conference Proceedings, IEEE International Conference on Vehicular Electronics and Safety, ICVES-2005, Xian, China, vol.1, pp 161–164
17. Hattori H et al (Sep. 7–10, 2009) Stereo-based pedestrian detection using multiple patterns. Research & Development Center, Toshiba Corporation, Conference Proceedings, BMVC-2009, London, UK, paper 243, pp 1–10
  18. Saxena A, Schulte J, Ng AY (Jan. 2007) Depth estimation using monocular and stereo cues. Stanford University, Conference Proceedings, 20th International Joint Conference on Artificial Intelligence, IJCAI-2007, Hyderabad, India, pp 2197–2203
  19. Trinh H, McAllester D (Sep. 7–10, 2009) Unsupervised learning of stereo vision with monocular depth cues. The Toyota Technological Institute, Conference Proceedings, BMVC-2009, London, UK, paper 432, pp 1–11
  20. Yokoyama E, Nagasawa M, Katayama T (1994) A disturbance suppression control system for car-mounted and portable optical disc drives. IEEE Trans Consum Electron 40(2):92–99
  21. Pan M-CH, Wei W-T (2006) Adaptive focusing control of DVD drives in vehicle systems. J Vibr Control 12:1239–1250
  22. Widrow B et al (1975) Adaptive noise cancellation: principles and applications. Proc IEEE 63(12):1692–1716
  23. Perrin B (Apr. 2007) The challenges of automotive vision systems design. Lattice Semiconductor Corp., White Paper, Hillsboro, Oregon, USA, pp 1–10
  24. Morris K (Mar. 30, 2004) FPGAs hit the road - programmable logic drives automotive applications. FPGA and Programmable Logic Journal, Design Article, FPGA and Structured ASIC Journal, [www.fpgajournal.com](http://www.fpgajournal.com), Portland, Oregon, USA
  25. Howell K (Feb. 27, 2007) Reprogrammable logic drives automotive vision systems design. Lattice Semiconductor Corp., Design Article, FPGA and Structured ASIC Journal, [www.fpgajournal.com](http://www.fpgajournal.com), Portland, Oregon, USA, pp 1–6
  26. Fraunhofer (Aug. 2007) GigE/Gigabit Ethernet standard investigation. Fraunhofer Institute for Photonic Microsystems (IPMS), Study Report, Dresden, Germany
  27. Sony Imaging, “Can GigE Vision deliver on its promise?”, Technical White Paper, available at: <http://www.sony-vision.com>, Sony Image Sensing Solutions, Accessed: Jan. 2008, Surrey, UK
  28. Pan J (Jan. 2003) Enhanced TCP/IP performance with AltiVec. Technical White paper, Freescale Semiconductor, Inc., Motorola Literature Distribution, Colorado, USA
  29. U J, Suter D (Nov. 2004) Using synchronised firewire cameras for multiple viewpoint digital video capture. Technical Report, Electrical and Computer Systems Engineering, Monash University, Clayton, Australia
  30. Muehlmann U, Ribo M, Lang P, Pinz A (Apr. 2004) A new high speed CMOS camera for real-time tracking applications. Graz University of Technology, Austria, Conference Proceedings, IEEE International Conference on Robotics and Automation, vol 5, New Orleans, LA, USA, pp 5195–5200
  31. Edens G, Hoover D, Meike R, Ryan T. Synchronous network for digital media. US Patent, No: US#6611537, Assignee Centillium Communications, Inc., Filed May 1998, Canada
  32. Hammerschmidt C (Apr. 2007) Inova, Fujitsu, BMW team for automotive multimedia bus. Design Article, EE Times Europe, Munich, Germany
  33. Römer M et al (Apr. 2009) Real-time camera link for driver assistance applications. In: *Advanced Microsystems for Automotive Applications 2009*, (Smart Systems for Safety, Sustainability, and Comfort), Springer Berlin Heidelberg, Book Chapter, pp 299–310, Berlin, Germany
  34. AIA (Jan. 2006) Camera Link® - Specifications of the Camera Link® Interface Standard for Digital Cameras and Frame Grabbers. Standard Specification Document, Version 1.2, AIA (Automated Imaging Association), Michigan, USA
  35. “Camera Link Fiber Extenders”, Product Brochures, Available at: <http://www.phrontier-tech.com>, Phrontier Technologies LLC, Accessed: Aug. 2008, California, USA
  36. Thinklogical, “Camera Fiber-Link”, Product Brochures, Available at: <http://www.thinklogical.com/product.asp?ID=32>, Thinklogical, Accessed: Jul. 2008, Connecticut, USA
  37. Maas HG (Jan. 2007) Concepts of single high-speed camera photogrammetric 3D measurement. Videometrics IX (IS&T/SPIE 19. Annual Symposium Electronic Imaging), Conference Proceedings, SPIE Proceedings Series Vol. 6491, San Jose, California, USA
  38. Kovacs J. “An Overview of Genlock”, Application Note No:5, MicroImage Video Systems, available at: <http://www.mivs.com/technical/appnotes/an005.html>, Accessed: Oct. 2001
  39. Trammell JE (Feb. 1986) Apparatus for synchronising two video pictures by controlling vertical synchronisation of a video camera. US Patent, No: US#4568976, Primary Classification: 348/516, New Jersey, USA
  40. Tashiro A (Apr. 1994) Method and apparatus for synchronising two cameras. US Patent, No: US#5307168, Primary Classification: 348/64, Tashiro, Atsushi, Sony Electronics inc (US), New Jersey, USA
  41. Chen, Tsuhan et al (Jan. 2002) Frame synchronisation in a multi-camera system. US Patent, No: US#6340991, Primary Classification: 348/513, AT&T Corp, Pennsylv., USA.
  42. Dexter E et al (Sep. 7–10, 2009) Multi-view Synchronization of Human Actions and Dynamic Scenes. INRIA, Vista Group, Campus Universitaire de Beaulieu, Conference Proceedings, BMVC-2009, London, UK, Paper 59, pp 1–11
  43. Trinkel et al (Mar. 2002) Synchronisation of a stereoscopic camera. German Patent, No: DE#10,044,032, Untermaubach, Germany
  44. Tserkovnyuk, Walter V et al (Sep. 2005) 3D camera. US Patent, No: US#6950121, Primary Classification: 348/47, Vrex Inc. (US), New York, USA
  45. Cooper AN et al (Nov. 1999) System and method for synchronisation of multiple video cameras. US Patent, No: US#5995140, Primary Classification: 348/159, Ultrax Inc (US), Texas, USA
  46. Cooper AN et al (Jan. 2004) Digital camera synchronisation. US Patent Application, No: US#20040017486, Primary Classification: 348/211.100, Texas, USA
  47. Matrox Imaging (Apr. 2007) Matrox Helios eCL/XCL. Datasheet/Brochure, Matrox Electronic Systems Ltd, Quebec, Canada
  48. Atmel Imaging (Mar. 2006) Automotive B&W VGA CMOS Image Sensor. AT76C410ABA Preliminary Datasheet, Rev. 0.5, Atmel Corporation, Grenoble, France
  49. Berger PD (Mar. 2006) CMOS High Dynamic Camera. Sensation, Characterisation Report, Atmel Corporation, Grenoble, France
  50. Mann S, Picard RW (May 1995) Being Undigital With Digital Cameras: Extending Dynamic Range by Combining Differently Exposed Pictures. Conference Proceedings, 48th IS&T's Annual Conference, Washington, DC, USA, pp. 422–428
  51. Debevec P, Malik J (Aug. 1997) Recovering High Dynamic Range Radiance Maps From Photographs. Conference Proceedings, ACM SIGGRAPH '97, Los Angeles, USA, pp. 369–378
  52. Sa A, Velho L (Feb. 2008) High dynamic range image reconstruction. Text Book, Series editor: Brian Barsky B., Morgan & Claypool Publishers, USA
  53. Helion-Blendfest GmbH “High Image Quality with HDR-CMOS Image Sensors”, Conference Proceedings, 4th Fraunhofer IMS CMOS-Imaging Workshop, May 2008, Germany