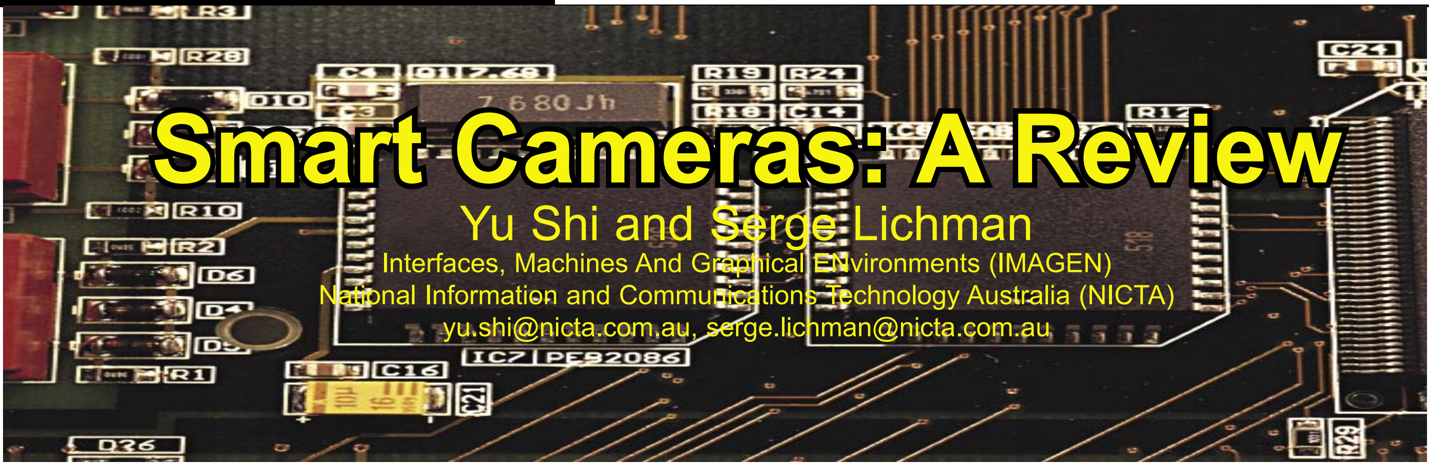


Smart Cameras: A Review

Yu Shi and Serge Lichman
Interfaces, Machines And Graphical ENvironments (IMAGEN)
National Information and Communications Technology Australia (NICTA)
yu.shi@nicta.com.au, serge.lichman@nicta.com.au

An interesting white paper about Smart Camera provided by Yu Shi and Serge Lichman (Part one)

Abstract

Smart cameras are cameras that can perform tasks far beyond simply taking photos and recording videos. Thanks to the purposely built-in intelligent image processing and pattern recognition algorithms, smart cameras can detect motion, measure objects, read vehicle number plates, and even recognize human behaviours. They are essential components to build active and automated control systems for many applications, and they will play significant role in our daily life in the near future. This paper aims to provide a first comprehensive review of smart camera technologies and applications. Here, we analyse the reasons behind the recent rapid growth of the smart cameras, discuss different categories of them and review their system architecture. We also examine their intelligent algorithms, features and applications. Finally we conclude with a discussion on design issues, challenges and future technological directions.

1 Introduction

What is a smart camera?

Different researchers and camera manufacturers offer different definitions. There does not seem to be a well-established and agreed-upon definition in either the video surveillance or machine vision industries, probably the two most active and advanced applications for smart cameras at present. For the purpose of this paper, *we define a smart camera as a vision system in which the primary function is to produce a high-level understanding of the imaged scene and generate application-specific data to be used in an autonomous and intelligent system.* ***The idea of smart cameras is to convert data to knowledge by processing information where it becomes available, and transmit only results that are at a higher level of abstraction.*** A smart camera is 'smart' because it performs *application specific information processing* (ASIP), the goal of which is usually not to provide better quality images for human viewing but to understand and describe what is happening in the images for the purpose of better decision-making in an automated control system. For example, a motion-triggered surveillance camera captures video of a scene, detects motion in the region of interest, and raises an alarm when the detected motion satisfies certain criteria. In this case, the ASIP is motion detection and alarm generation.

The important differences between a smart camera and "normal" cameras, such as consumer digital cameras and camcorders, lie in two aspects. The first is in camera system architecture. A smart camera usually has a special image processing unit containing one or more high performance microprocessors to run intelligent ASIP algorithms, in which the primary objective is not to improve images quality but to extract information and knowledge from images. The image processing hardware in normal cameras is usually simpler and less powerful with the main aim being to achieve good visual image quality. The other main difference is in the primary camera output. A smart camera outputs either the features extracted from the captured images or a high-level description of the scene, which is fed into an automated control system, while for normal cameras the primary output is the processed version of the captured images for human consumption. For this reason, normal video cameras have large output bandwidth requirements (in direct proportion to the resolution of the image sensor used), while smart camera can have very low data bandwidth requirements at the output (it can be just one bit in the simplest case, with '1' meaning 'there is motion' and '0' meaning 'there is no motion',

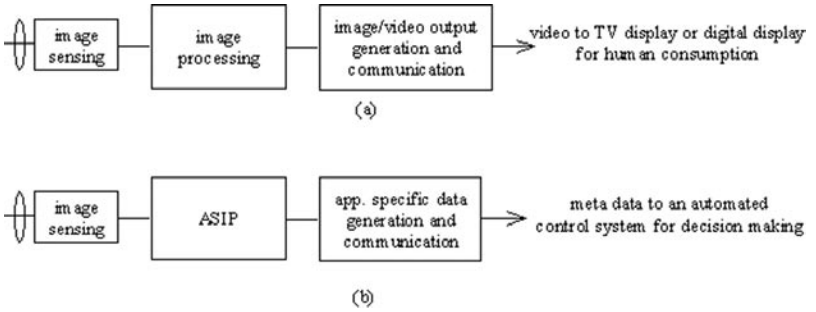


Figure 1: Differences between a normal camera (a) and a smart camera (b).

for example). These differences are illustrated in figure 1.

Smart cameras can exist where a camera is not expected to be. A good example is the ubiquitous optical mouse for PC. Most optical mice contain a miniature digital video camera inside the mouse casing. They work by shining a bright light onto the surface below, then using a camera to take up to 1500 pictures a second of that surface. An intelligent image processing circuit inside the mouse performs image enhancement and calculates the mouse motion based on image difference between successive frames. This difference is then used to displace the mouse cursor on the screen. The optical mouse is a good

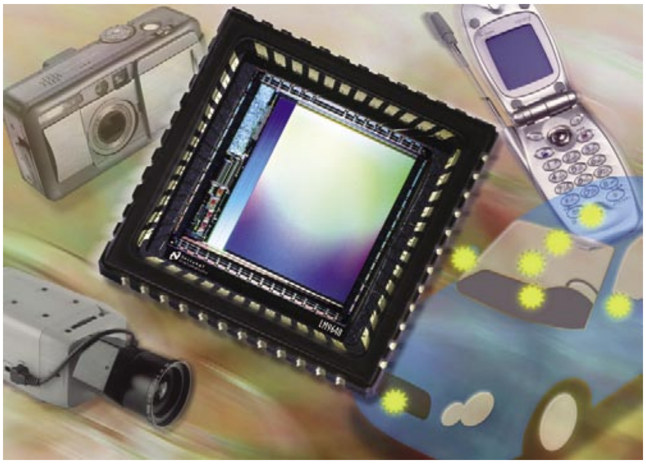
example of a smart camera in three respects: firstly it's a stand-alone camera with camera and processing in a single embedded device; secondly the camera is used not to take pictures or video for human consumption, but to produce a feature vector (motion vector of in x and y directions) to represent the object (the mouse in this case) displacement; thirdly it shows that smart cameras are not restricted to a niche market, but can be adopted ubiquitously.

Strictly speaking, a smart camera is a stand-alone, self-contained device that integrates image sensing, ASIP and communications in one single box. It is designed for a special type of application (for example, surveillance, and industrial machine vision). However, there are other types of vision systems that are often referred to as smart cameras as well, such as PC-based smart cameras. We'll analyse these different types of smart cameras in section 3. The term 'smart camera' in this paper covers both stand-alone smart cameras and other types of smart cameras, as described in section 3.1, unless specified otherwise.

One of the smart projects at Siemens

The advent of smart cameras can be traced back to the early 1990s when PCs became popular and video frame grabbers became available. Early solid state CCD (Charge-Coupled Device) cameras of the mid-1970s were analog cameras. Later digital signal processing (DSP) technologies pushed analog CCD cameras into the digital era with enhanced image quality, but the output of most of these cameras was still analog (e.g. NTSC/PAL signals). Frame grabbers allowed CCD cameras with analog output to be connected to computers and digitized for versatile processing by computers. This marked the beginning of smart camera systems, with the camera performing image capture and computer carrying out intelligent processing tasks such as motion detection and shape recognition. The first applications were in the area of industrial machine vision and surveillance.

The real interest in and the growth of smart cameras started in late 1990s and early 2000s, spurred by factors such as technological advancements in chip manufacturing, embedded system design, the coming-of-age of CMOS (Complementary Metal Oxide Semiconductor) image sensors and so on. Market demands from surveillance and machine vision also played significant roles. Advanced smart camera systems often integrate the latest technologies in image sensors, optics, imaging systems, embedded systems, computer vision, video analysis and communication, networking and etc.



A CMOS chip (photo curtesy of National Semiconductor)

The heart of smart cameras is the intelligent ASIP algorithms and the hardware that runs them. Image feature extraction and pattern recognition are probably among the most widely used algorithms in smart cameras. In a way, a smart camera can be thought of as an image feature extractor or a visual pattern recogniser. Research in computer vision, image understanding and pattern recognition has yielded many algorithms and solutions that can be used by smart cameras. However, the performance and robustness of the ASIP algorithms when deployed into cameras operating under real-world conditions are among the most important issues facing the development and commercialization of new smart cameras.

In the remainder of this paper, we analyse the main reasons behind the rapid growth of smart cameras (section 2), review system architecture of different smart cameras (section 3), review the state-of-the-art smart camera systems and ASIP algorithms for some applications (section 4), and finally discuss some design issues and conclude with some thoughts about technical challenges and future technological directions (section 5).

2 The Rapid Growth of Smart Cameras

2.1 Coming of Age of CMOS Image Sensors

The advent of CMOS image sensors (CIS) in late 1990s played an important role in the development of smart camera technology and systems, and has potential to make smart camera smaller, cheaper and more pervasive. Compared to CCD, CIS have several advantages which make them excellent candidates for smart camera front-end. These include smaller size, cheaper manufacturing cost, lower power consumption, the ability to build a camera-on-a-chip, the ability to integrate intelligent processing circuits onto the sensor chip, and significantly simplified camera system design.

Most CIS's are manufactured using the same process by which semiconductor chips (CPUs, memories, etc) are made. This means that many semiconductor manufacturers can make CIS, which drives up competition and reduces cost. CCD sensors, by contrast, are made using special chip manufacturing process and there are only a few manufacturers in the world, mostly in Japan. CCD-based camera chip-sets usually include at least three or four chips: a CCD pixel array, CDS (Correlated Double Sampling), a timing generator, and ADC (Analog-to-Digital Converter). In the case of CIS, all these functions can be integrated onto one single chip, making it a real camera-on-a-chip with light in and pixel out. This greatly simplifies camera system design and reduces cost. Compared with the CCD chip-set, there are many more sources from which a CIS can be purchased, even a single item at a time, which is very difficult to achieve in the case of CCD. All this makes it much easier for more researchers, students, and camera manufacturers alike to develop smart cameras of their own.

Probably the most important advantage of CIS over CCD lies in its ability to have image sensor array and intelligent image processing circuits side by side on the same chip. This makes a single chip smart camera possible. One example is a vision-based single-chip fingerprint reader with on-chip CIS, a processing circuitry performing pattern matching and a memory storing templates of one or several user fingerprints for real-time comparison and identification [1].

A recent market survey by Gartner Dataquest [2] estimated that there are about 40 suppliers of CIS world-wide, and that the global CIS market would increase from $3.2 billion in 2005 to 5.6 billion by 2008. The survey showed that automobile, medical imaging and surveillance applications are among the emerging markets for CIS products.

IC Machine analysis with video camera (courtesy of Microscan)

2.2 Research in Computer Vision and Pattern Recognition

What makes a camera smart is the intelligent ASIP - the application-specific information processor built into the camera system. The advancement in academic and industrial research in real-time image processing and understanding, pattern recognition, machine learning, computer vision and video communication continues to provide a large library of intelligent algorithms for use by smart cameras for different applications. As an example, Intel's OpenCV (Open Source Computer Vision) Library [3] has been very popular with academic researchers and students working on smart camera projects. Every year, numerous international journals, conferences and workshops give researchers world-wide forums to present their innovative work in areas such as computer vision and pattern recognition. A lot of the work presented can be seen as embryos of future smart cameras. Recently, first ever international conferences and workshops have been held focusing on the design of embedded vision systems.

2.3 Embedded System Technologies

A stand-alone smart camera is essentially an embedded vision system. Compared with PC-based systems, an embedded system is usually subject to many constraints on the design, implementation and production of the device which encapsulates it, such as low power, limited resources, real-time processing and low cost. An embedded vision system is even more challenging to design due to video processing's insatiable demand for computing power and memory resources. In the last decade, embedded vision systems have made great progress thanks to the increasing affordability of powerful processors and memory chips, availability of real-time operating systems, low complexity intelligent algorithms and the coming-of-age of system development software and tools.

Functional integration seems to be a trend in consumer electronics and ICT (Information and Communications Technology). For example, many cellular phones now come with a camera and can play music and receive radio. Some webcams have built-in intelligence such as face tracking. Functional integration can seemingly make a normal camera become smart. For example, a camera with an integrated voice/sound detection component can take a picture of the surrounding area when a human voice is detected, or it can take a picture in a direction from which a gun-shot has been detected [4].

2.4 Socio-Economical Drivers

Thanks to Moore's law, semiconductor chips and computer hardware continue to shrink in size, reduce in cost and gain in performance. This has driven the prices of cameras, frame grabbers and computers down and made smart camera systems, especially PC-based systems, more affordable to research and development on one hand and to the market and end-users on the other. As hardware constraints (cost-wise) are lifted, software developers have more freedom to write "smarter" algorithms.

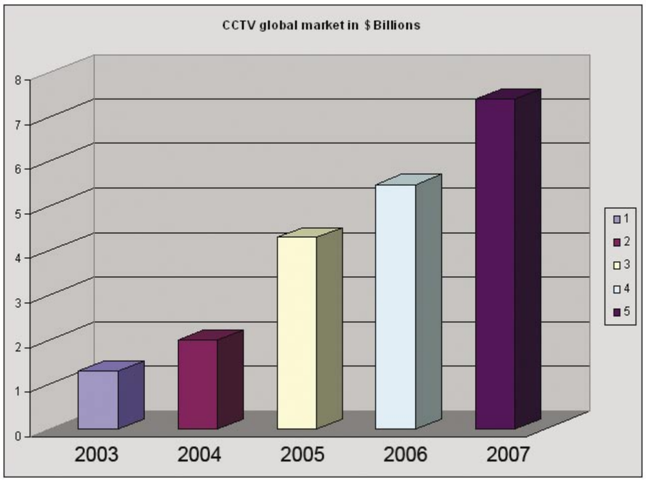
One of the most significant developments in surveillance and security industries in the last several years has been the wide use of CCTV (Closed Circuit Television) cameras and their impact on crime, terrorist attacks, and on the general public. It is noticeable that after the 9/11 event in the US, video surveillance has received more attention not only from the academic community, but also from industry and governments.

The recent terrorist attacks in the London Underground in mid-2005 and the successful use of CCTV by police in identification of perpetrators have intensified the talk about a new generation of intelligent video surveillance systems based on smart cameras. In fact, surveillance and security demands are an important driving force behind the ever-increasing scale of academic and industrial research in advanced vision algorithms such as object tracking and identification, and human behaviour analysis.

2.5 Market Demands and Analysis

2.5.1 Digital Video Surveillance

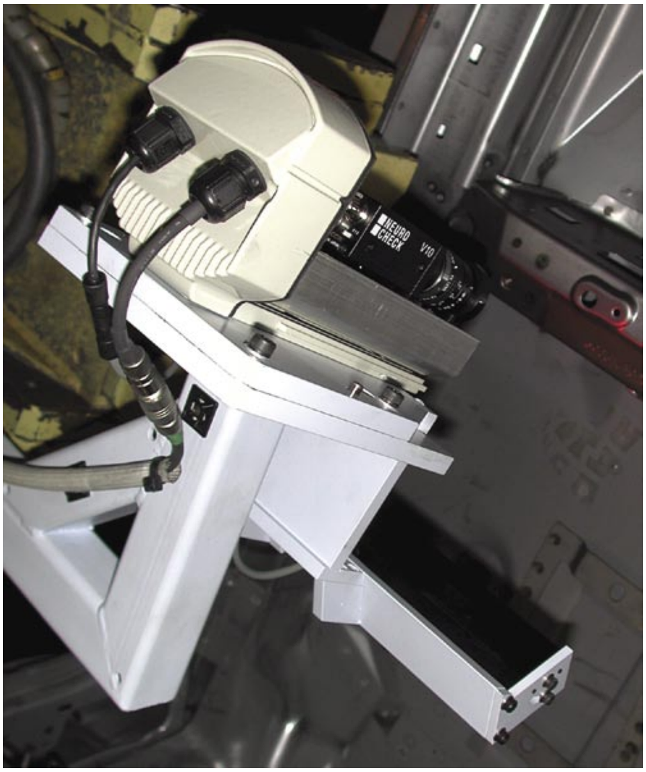
The first generation of CCTV cameras (1980s-1990s) was mostly analog cameras with limited functionality and high cost. Digital CCTV cameras and the use of DVR (Digital Video Recorders) represented the second generation (2G, 1990s-now). Digital CCTV cameras built using CCD and CMOS image sensors provide better video quality, some intelligent functions such as motion detection, electronic PTZ (Pan-Tilt-Zooming), and networking. The 2G CCTV systems have become mass market products, fuelled by improved affordability and society's increasing concerns over safety and security. According to estimates made in 2004 by market research firm Datamonitor [5], digital video surveillance is a high-growth segment within the overall surveillance market estimated at 55% CAGR (Compound Annual Growth Rate) between 2003 and 2007. In dollar terms, between 2003 and 2007 the market will grow from US$1.3bn to US$7.4bn globally.



However, the 2G CCTV systems are not "smart" enough to help prevent crimes or terror attacks, even though they proved very useful in post-event identification of crime perpetrators. The 2G CCTV systems are mostly not automated systems and rely strongly on trained security personnel to perform image analysis, object tracking and identification. The increasing number of cameras makes this difficult for real-time analysis by security personnel. Network bandwidth is another important issue affecting real-time processing needed for crime prevention. The intelligent video surveillance system (IVSS) (also called the third generation CCTV system) will try to provide solutions to these problems. Smart cameras will be one of the fundamental building blocks of the IVSS, making it possible to build and deploy automated, distributed and intelligent multi-sensory surveillance systems capable of tracking humans and suspected objects, analysing human behaviours, and etc. Many market research firms have predicted significant growth in intelligent video systems and smart cameras. For example, the market researcher Frost & Sullivan [6] has forecast that the US$153.7 million video surveillance software market is expected to witness a healthy CAGR of 23.4% from 2004 to 2011 to reach US$670.7 million.

2.5.2 Industry Machine Vision

Industrial machine vision is probably the birth place of smart cameras, at least in terms of the systematic use of commercial smart cameras. It is also one of their most active playgrounds. Most machine vision smart cameras are stand-alone cameras. The demand for these cameras has been steadily increasing over the years. The major end user industries are in robotics, semiconductor, electronics, pharmaceutical, manufacturing, food, plastics and printing. The tasks these smart cameras usually perform include bar-code reading, part inspection, flaw detection, surface inspection, dimensional measurement, assembly verification, print verification, object sorting, OCR (optical character recognition) and maintenance. A recent survey on machine vision products from a Europe based market research firm IMS Research [7] has discovered that smart cameras are rapidly accounting for a greater share of the machine vision market revenue. Demand for smart cameras is primarily driven by the increasing demand for better production efficiency and quality control in industries such as manufacturing and medicine / pharmaceutical. The survey revealed that whilst the sale of more traditional PC-based products (cameras and frame grabbers) has fallen, sales of smart cameras and compact vision systems have continued to grow. The survey predicts that the machine vision market in Europe will grow at an average rate of 11.6% each year to 2006. The highest levels of growth, approaching 20%, are



forecast for the smart sensor and cameras product groups resulting in more than doubling in value in dollar terms. The same trend has also been forecast by the same company for the Asia-Pacific market [8]. An estimate provided by the annual market study by the AIA (Automated Imaging Association) for the 2003 North American machine vision smart camera market is about $57 million US dollars, with growth at 15% per year in terms of revenues and 20% per year in terms of units [9].

2.5.3 Other Significant Markets

Other important markets for smart cameras are ITS (Intelligent Transport Systems), automobiles, HCI (Human Computer Interface), medical/healthcare, games, toys, video conferencing, biometrics.

3 Review of Smart Camera System Architecture

In recent years, smart cameras have attracted considerable attention from academic and industrial research and development (R&D) organizations. However, to the best of the authors' knowledge, a systematic approach to analysing smart cameras has yet to be agreed-upon. In this section we firstly present one approach to classify smart camera systems and provide an analysis of their system architecture, followed by a review of some R&D activities on the design of smart cameras as embedded systems.

3.1 Classification of Smart Cameras

Smart cameras can come in different system and physical configurations. Figure 2 shows one proposed classification of different types of vision systems and smart cameras.

As shown in Figure 2, stand-alone smart cameras are a subset of embedded vision systems. Non-stand-alone embedded smart cameras are sometimes called compact vision systems. Compact vision systems are usually composed of general purpose cameras connected to an external embedded processing unit in a separate box to provide ASIP and communication/networking functionality. Single-chip smart cameras can be thought of as a special case of smart cameras because they require special system design considerations and are usually used in carefully targeted applications. Non-stand-alone smart cameras can be thought of as virtual smart cameras because from user point of view the cameras are smart, even though the ASIP which makes them smart may be performed by an external unit, like a hardware accelerator board, a local PC or a networked PC. PC-based smart cameras, consisting of a general purpose video camera, a frame-grabber of some sort and a PC, of which the CPU performs the ASIP, is a very common and inexpensive platform for researchers, academics and students to conduct research on smart cameras. Sometimes a normal camera is connected to a PCI (Peripheral Component Interconnect) processing board within a PC. In this case, the PCI board may perform most of the ASIP and output generation, while the PC provides a flexible operator interface or additional processing power. This kind of system is a special case of a compact vision system and a PC-

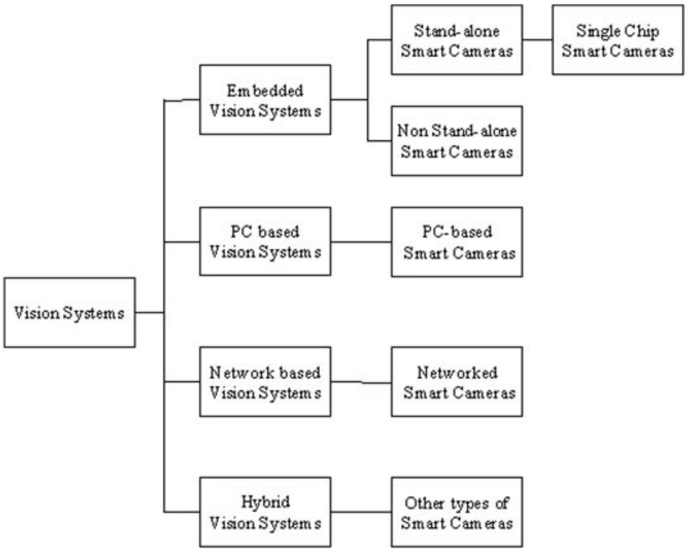


Figure 2: One proposed classification of vision systems and smart cameras.

based system. A digital CCTV surveillance system with intelligent features is an example of a network-based smart camera system, and the next generation of distributed intelligent video surveillance systems will be the exciting test ground for smart cameras, especially stand-alone smart cameras. Hybrid vision systems may give rise to some special types of smart cameras. This category may also include smart camera systems that may need some kind of human intervention to help provide high accuracy data output.

3.2 Analysis of Different Types of Smart Cameras

3.2.1 Common Characteristics

The common basic components of a normal digital video camera (consumer, professional or industrial) include optics, solid-state image sensor (CCD or CMOS), image processor(s) and supporting hardware, output generator, and communication ports. The main tasks performed by the image processor(s) are to provide colour interpolation, colour correction or saturation, gamma correction, image enhancement and camera control such as white balance and exposure control. The output generator can be an NTSC/PAL encoder to provide standard TV-compatible output, or a video compression engine to provide compressed video streams for communication over network, or digital video output generator such as a Firewire encoder. Communication ports, such as Ethernet or RS232 provide the basis for networked camera functionality or camera configuration and firmware upgrading through a PC respectively.

The main basic components of a smart camera typically exhibit all the above essential components of a normal camera, with the following differences:

- A smart camera has a distinct and powerful signal processing unit to perform image feature extraction and/or pattern analysis based on application-specific requirements; and
- A smart camera has an output generator to produce a coded representation of the image features and/or results from the pattern matching, or in some cases, control signals for other devices (e.g. alarm triggering signal) or actions (e.g. sending a picture of the number plate of a car which is speeding to police).

System architecture design for smart cameras often involves significant system engineering effort. Clear application requirements and specifications are crucial to the successful design. Software architecture, hardware architecture, and network architecture for network-based systems, need to be jointly designed to maximize resource usage and efficiency, and to reduce cost and time-to-completion. More detailed design considerations are discussed in section 5.1.

3.2.2 Stand-alone Smart Cameras

A stand-alone smart camera integrates image capture, ASIP and application specific output generation into a single device casing. A stand-alone smart camera may look very much like a normal industrial camera or a CCTV camera. While the primary function of a normal camera is to provide raw video for monitoring and recording, a smart camera is usually designed to perform specific, repetitive, high-speed and high-accuracy tasks in industries such as machine vision and surveillance. Most of the industry machine vision cameras are stand-alone smart cameras. While a normal video camera may only cost anywhere between US$50 and US$2 000, a machine vision smart camera can cost between US$1 000 and $6 000 per unit [10] and beyond, depending on the functionality and level of customization.

Many pattern recognition techniques involve two types of processing tasks, data-intensive tasks such as image enhancement and feature extraction, and math-intensive tasks such as statistical pattern matching. While data-intensive tasks require high speed hardware to deal with high pixel volume and high frame rate, math-intensive tasks often require high performance processors to deal with issues such as pipelining and floating-point arithmetic. For demanding applications, camera hardware architecture may be based on a heterogeneous- and multiple-processor

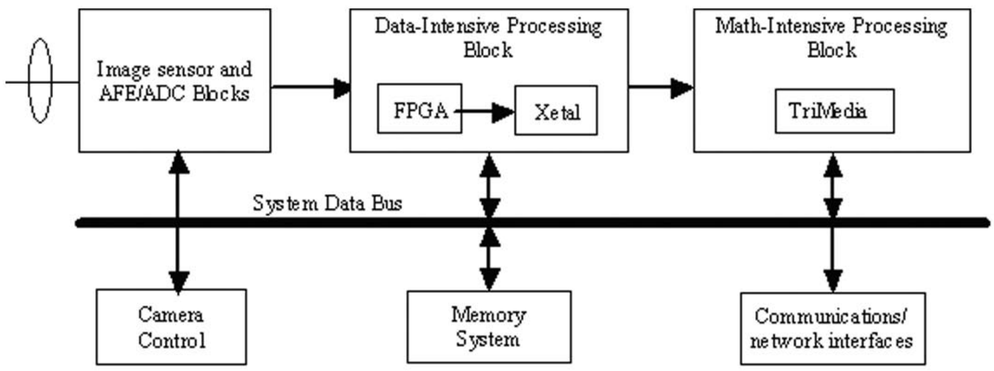


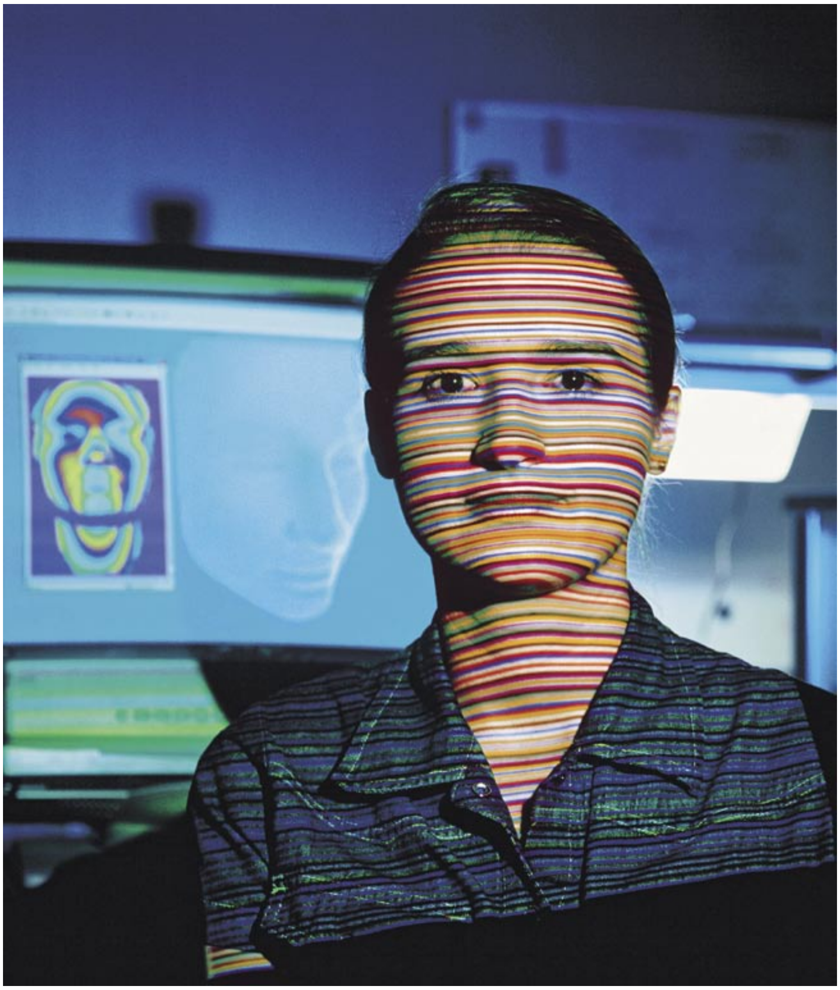
Figure 3: A stand-alone smart camera system architecture for face recognition

platform, with one or more processor(s) capable of implementing parallel processing (e.g. an FPGA - Field Programmable Gate Array) performing data intensive tasks, and a DSP and/or a RISC (Reduced Instruction Set Computer) processor performing math-intensive tasks. A smart camera built for face detection and recognition application by Broers et al. [11] is such an example. The system employs an FPGA and a parallel processor Xetal working in SIMD (Single Instruction Multiple Data) mode, to perform data intensive operations such as face detection. A high performance DSP, TriMedia, with a VLIW (Very Long Instruction Word) core is used to perform high level programs such as face recognition. The system architecture can be represented as in Figure 3.

3.2.3 Single-Chip Smart Cameras

Single-board or single-chip smart cameras are a special kind of stand-alone smart camera. Single chip smart cameras take advantage of the integration capability of CMOS image sensors by building intelligent ASIP circuits onto the image sensor chip, potentially releasing the host computer of cumbersome pixel processing tasks and minimizing the data transfer between camera and computer. In some cases, pixel-level ADC and processing can be achieved [12], which can lead to a brand new level of signal and image processing methodologies. Single-chip smart cameras make it possible to design very efficient, very small, low power and low cost cameras (when a large volume is produced). As examples, the VISoc single chip smart camera [13] integrates a 320x256 pixel CMOS image sensor, a RISC processor, a vision co-processor and I/O onto a single chip, which has been fabricated in a 0.35um process on an area of about 36mm^2, and a typical power dissipation of about 1W at 3.3V at 60MHz. Moorhead et al. [14] designed a smart CMOS camera chip which integrates an edge detection mechanism directly into the sensor array. Lee et al. [15] reported the design of a 30 frames/second VGA-format CMOS image sensor with an embedded massively parallel processor, for real-time skin-tone detection.

In some applications single chip smart camera can bring distinct advantages. For example, Shigematsu et al. argue that, compared with conventional multi-chip fingerprint readers, a single-chip smart camera based fingerprint reader can have advantages of being much smaller, allowing much simplified integration into mobile devices such as mobile phone, being low in cost, and having improved security [1]. The main disadvantage of the single-chip smart camera lies in the cost of chip design and manufacturing, unless a large volume of units can be produced to justify the initial capital investment. Nevertheless, a single-chip smart camera is a smart sensor that has potential to make vision systems pervasive, especially when connected to wireless sensor networks.



3D Face analysis with video (courtesy of Siemens)

3.2.4 Embedded System based Smart Cameras

This category of smart cameras most often consists of a camera (usually a general purpose one) and an external embedded processing unit connected to it. For example, an embedded system based smart camera could be a general-purpose camera connected to a high performance video processing board, which itself is connected to a PC, either through a PCI slot or through a RS232 port. This kind of configuration is not too different from a PC-based system. Many 2G digital CCTV systems with some intelligent features belong to this category.

The necessity of having a dedicated and embedded processing unit in this type of smart cameras is due to the fact that PC, while flexible and versatile, is far from being adequate to perform intensive image and video processing and pattern recognition tasks, particularly when high-resolution, high frame rate and low latency processing is required. Another advantage of this kind of system is that once proof-of-concept is achieved and end-users are identified, it is easier for the system to be converted to a stand-alone smart camera if required.

Smart cameras used in robotic and automobile applications can also be classified into this category. These cameras may share computing resources such as a processor and memory with other embedded devices in the robot and in the vehicle.

3.2.5 PC and Network based Smart Cameras

PC-based smart camera systems are probably most popular within the academic research environment, as a first step to conducting computer vision and pattern recognition research, and building first prototype for proof-of-concepts. It is a very simple and inexpensive configuration, as the prices for general purpose video cameras and PCs continue to fall. Most often, a general purpose camera is connected to a PC through either a frame grabber or a communication port such as USB, Firewire, CameraLink, or Ethernet. This type of system relies on the PC's CPU to perform image analysis, feature extraction and pattern recognition tasks. The availability of various vision processing libraries for PC platforms makes this kind of system very popular. PCs also provide a more flexible environment for building user interfaces.

USB cameras, Firewire cameras and network cameras allow digital images to be transferred directly from camera to a PC or an embedded processing hardware, avoiding signal integrity loss caused by DAC (digital to analog conversion) inside many CCTV cameras and ADC by frame grabbers. For high-resolution cameras, Firewire cameras are starting to become popular and affordable, but CameraLink remains dominant, especially for high bandwidth and high performance applications.

The 2G CCTV system is a network based video surveillance system (NVSS). An NVSS with built-in intelligent surveillance features can be loosely considered as a network of virtual smart cameras. An NVSS is composed of four main layers: a CCTV camera (sensor) layer, a network layer, a central computer layer and a trained security personnel layer (Figure 4). As discussed in section 2.5.1, in most of the currently deployed NVSSes, the ASIP tasks such as object tracking and identification and threat detection are typically performed mostly by trained security personnel. However, human monitoring of surveillance video is a very labour-intensive task. It is generally agreed that watching video feeds requires a higher level of visual attention than most every day tasks. Specifically vigilance, the ability to hold attention and to react to rarely occurring events, is extremely demanding and prone to error due to lapses in attention. A recent study by the US National Institute of Justice found that, after only 20 minutes of watching and evaluating monitor screens, the attention of most individuals will degenerate to well below acceptable levels [16]. The next generation of video surveillance systems - intelligent video surveillance systems (IVSS) – will try to solve these problems by providing automated video surveillance and crime pre-emption abilities. The IVSS will seek a re-distribution of ASIP tasks among the four layers in the NVSS system, notably shifting processing load from security personnel to central computers or DVR (in short-term), and probably more importantly to the surveillance cameras – that is, the introduction of (stand-alone) smart cameras to replace passive or dumb CCTV cameras (in mid- and long-term). The use of smart cameras would greatly reduce the bandwidth problem caused by the increasing number of cameras present in the system and enhance surveillance system performance, as sending raw pixels over the network is less efficient than sending the results of intermediate analysis results. Smart cameras can also help in decentralizing the overall surveillance system, which can lead to improved fault tolerance and the realization of more surveillance tasks than with traditional cameras [17].

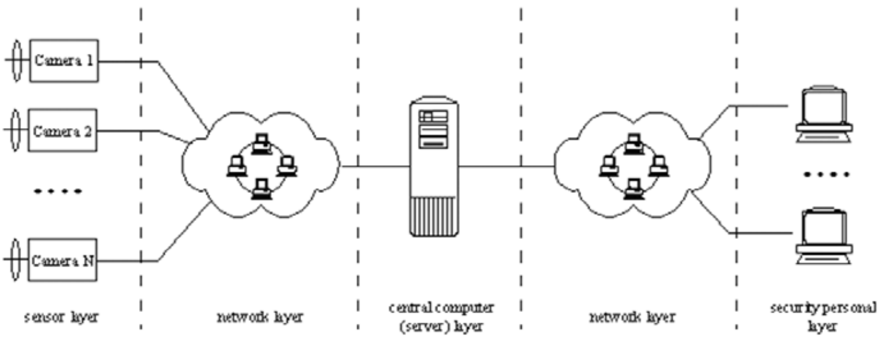
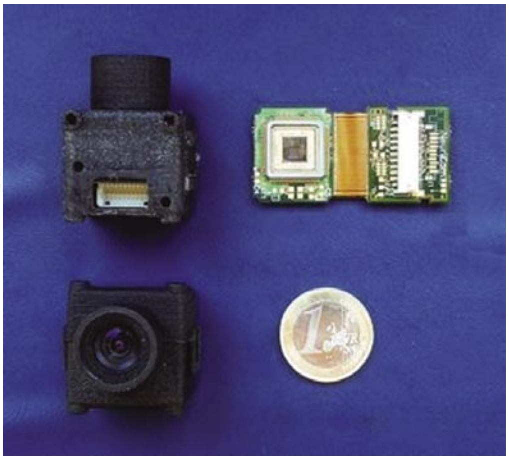


Figure 4: Four layers of a network based video surveillance system (NVSS).

3.3 Research in Smart Cameras as Embedded Systems

Video processing is notoriously hungry for computation horsepower, memory and other resources. Smart cameras as embedded systems have to meet the insatiable demand of video processing on one hand, and to meet the challenging demands of embedded systems, such as real-time, robustness, reliability under real-world conditions, on the other hand. This has made smart cameras a leading-edge application for embedded systems research [18]. Recently there has been a significant increase in research in building smart cameras as embedded systems. The first IEEE workshop on Embedded Computer Vision (ECV'05) was held in June 2005 [19]. The workshop addressed issues such as how to design smart algorithms to efficiently utilize embedded hardware, how to meet real-time constraints in embedded environment and verification methods for mission-critical embedded vision systems. In particular, the workshop discussed the suitability of FPGA for embedded vision systems.



Apart from numerous research groups working on developing smart cameras for video surveillance, there are a number of academic research groups in the world dedicated to research into building smart cameras as embedded systems. One prominent group is the Embedded Systems Group in Princeton University's Department of Electrical Engineering [18]. This group has developed an embedded smart camera system that can detect people and analyse their movement in real time. They are also working on a VLSI (Very Large Scale Integration) smart camera. An interesting research activity involving the design of stand-alone smart cameras is the SmartCam project at University of Technology Eindhoven [20]. This project investigates multi-processor based smart camera system architecture and addresses the critical issue of determining correct camera architectural parameters for a given application domain. Another project bearing the same name is being undertaken by the University of Technology in Graz, Austria [17]. The project aims to develop distributed smart cameras for traffic surveillance applications. They also investigate various issues involved in making smart cameras as embedded systems, such as resource-aware dynamic task allocation systems to support real-time requirements.

Many industry research groups and companies are involved in smart camera research for machine vision, especially in Germany, Japan and the US. There exist some very informative and useful journals and web portals for the machine vision world, such as IEEE Transactions on Pattern Analysis and Machine Intelligence, Advanced Imaging Magazine [21], Machine Vision Resources [22], Machine Vision Online [23].

A search on USPTO (US Patent and Trademark Office) web site can reveal many patents filed or issued in relation to the concept and embodiment of smart cameras as embedded systems. For example, patent #6 985 780 filed in Aug 2004 under the title of "Smart Camera" [24] made claims about a camera system that includes an image sensor and a processing module at the imaging location that processes the capture images prior to sending the results to a host computer. The processing module can perform tasks such as image feature extraction and filtering, convolution and de-convolution methods, correction of parallax and perspective image error and image compression.

About the Authors – YU SHI is a Senior Researcher with National ICT Australia in Sydney, Australia. He was granted his B.Eng in 1982 by the National University of Defense Technology in Changsha, Hunan, China. He later obtained his M.Eng and PhD in signal processing and biomedical engineering in 1988 and 1992 respectively in Toulouse, France. He also completed post-doctoral research at Oxford Brookes University in England in the late 1990s. His main research interests are in embedded vision systems, FPGA-based design and applications, multimodal user interfaces and web services.

SERGE LICHMAN is a Senior Research Engineer with National ICT Australia in Sydney, Australia. He received M.Eng in Electrical Engineering in 1988 by the Odessa State Polytechnic University in Ukraine. His 12 years of experience in the area of image and signal processing for commercial software and hardware gave him practical skills in full product development life cycles, from research to deployment. His work has led to several publications.

In the next issue - ASIP and Algorithms, Design considerations, etc.