A Future for Integrated Diagnostic Helping

Mathieu Thevenin, Anthony Kolar

A Future for Integrated Diagnostic Helping

Mathieu Thevenin and Anthony Kolar
CEA, LIST, Embedded Computing Laboratory
France

1. Introduction

Medical systems used for exploration or diagnostic helping impose high applicative constraints such as real time image acquisition and displaying. This is especially the case when they are used in surgical room where a high reactivity is required from operators. Large computing capacity is required in order to obtain valuable results. Integrators mainly prefer the use of general purpose architectures such as workstations (Gomes, 2011). They have to cope with manufacturing cost and setup simplicity. As general purpose devices need a large amount of space, the main part of the processing is deported from the handled diagnostic tools to an external unit. For example, this is the case of endoscopic device. Today, dedicated rooms are usually used for this purpose in many hospitals. Their associated external computers that are used for diagnostic system are cumbersome and are also energy consumers. These issues are too problematic to use efficiently these systems in a limited space. Indeed, they restrain the movements of the medical staff and complexify the deployment on the ground for military or humanitarian operations. Therefore it seems logical to integrate the maximum computing capacities diagnostic into helping devices themselves to make them completely handleable.

A large part of computing requirement of these systems is devoted to image processing. They can be quite simple like images reconstruction and enhancement, features detector or 3D reconstruction. Today, a large part of these processing is mainly embedded inside handled consumer's devices such as digital cameras or advanced driving assistance systems (ADAS). By the analysis of both medical and consumer's applications systems, it is possible to notice that they rely on similar algorithmic approaches. Also, most of integration constraints are similar if someone wants to miniaturize these consumer devices. This mainly concerns the chips silicon areas, their power consumption and their computing capacities. For example, a digital video sensor and image processor integrated to a cell phone cannot reach more than a half watt of power consumption for a silicon area of less than a dozen square millimeters. This is also the case for one of the most integrated medical diagnostic device which is the endocapsule. It form factor (Harada, 2008) limits components size while its autonomy is driven by energies efficiency. The whole device may not exceed a Watt of power consumption. About a half Watt is devoted to the part dedicated to computation for diagnostic, especially based on image processing. However, this part depends on the device features, such as communication systems and mechanical elements that may be used for mobility or biopsy.

Integrators also demands versatility in order to design unique products that can be used for different targets. For example, endoscopic exploration of larynx or intestinal and lung exploration do not uses the same devices, but these applications are all based on similar

image processing with minor variations. Moreover, these systems should be updatable to follow the science developments.

These requirements are also valid for large market devices such as cell phones and cameras. For example, general purposes or specific embedded processors are widely used like ARM microprocessors and Texas Instrument Digital Signal Processors (DSP) which are integrated into transportation, photonics, communications or entertainment (Texas Instrument, 2006). These markets drive both academic and industrial researches. The background knowledge is present inside laboratories; however its transfer to medical applications is not yet completely industrially ready.

This chapter provides clues to transfer consumers computing architecture approaches to the benefit of medical applications. The goal is to obtain fully integrated devices from diagnostic helping to autonomous lab on chip while taking into account medical domain specific constraints.

This expertise is structured as follows: the first part analyzes vision based medical applications in order to extract essentials processing blocks and to show the similarities between consumer's and medical vision based applications. The second part is devoted to the determination of elementary operators which are mostly needed in both domains. Computing capacities that are required by these operators and applications are compared to the state-of-the-art architectures in order to define an efficient algorithm-architecture adequation. Finally this part demonstrates that it's possible to use highly constrained computing architectures designed for consumers handled devices in application to medical domain. This is based on the example of a high definition (HD) video processing architecture designed to be integrated into smart phone or highly embedded components.

This expertise paves the way for the industrialisation of intergraded autonomous diagnostic helping devices, by showing the feasibility of such systems. Their future use would also free the medical staff from many logistical constraints due the deployment of today's cumbersome systems.

2. Video processing in diagnosis helping system

Since many years, the research about diagnosis helping devices is very active. This is true in both academic and industrial world. This can be explained by the fact that the possibilities of data analysis systems are becoming more and more complex and can extract a large amount of information. The helping in diagnosis can provide a solution to decrease the response time of the practitioner in urgent case or to help him in the preparation of the patient operation. This section firstly presents endocapsules as an example of one of the most constrained integrated diagnostic devices. It also representative of some of the major research domains in biomedical technology: image and signal processing, robotics and in-vivo communication. Next, the needs for diagnosis helping for such devices are presented, followed by an introduction about low level image processing in both consumers and medical components. Finally similarities between these two applications are developed and clues are given to appreciate required capacity for these embedded algorithms.

2.1 Researches in diagnostic helping devices
Researches in diagnostic helping devices cover a large number of domains, however one can focus on three items that emphasize the conception of an autonomous device. This is illustrated by Lab-On-Chips projects (Harada, 2008) that are able to do an auto diagnostic.

1. The video processing:

In the case of endocapsule (Karargyris, 2010), the main goal of video processing is to analyse a video sequence in order to find different features like bleeding, polyps, tumours, etc. Theses kinds of diagnosis helping are usually done in two steps. First a camera equipped device grab images to diagnostic, these images are then transmitted through a wireless connection to a workstation that analyzes them during an off-line processing. Figure 2.1 depicts the PillCam by GivenImaging, and the Endocam by Olympus can also be cited;

Fig. 2.1. PillCam by GivenImaging

2. The mechanical systems for autonomous devices:

Some researches focus on the integration of mechanicals devices to endocapsules in order to give them the ability to surgery using micro-instrumentation such as biopsy. An example of such an endocapsule is "Miro" (Kim and al., 2007) under Korea's Frontier 21 project as shown on Figure 2.2. The "Scuola Superiore Sant'Anna" (Quirini, 2007) also tries to integrate small mechanic legs to a video-capsule in order to give the practitioner the ability to move freely in the intestinal system.

Fig. 2.2. Principe of the endocapsule ''Miro'' and prototypes of a mobile endo-capsule

3. The communication and transfer protocol:

Communication protocol in human body is defined by the norm IEEE 820.15. Its frequency is 403 MHz. This is defined by the norm for in-vivo electronic devices. Antennas for this band are small while low emitting power is required due to limited loss of the signal in the environment. Moreover this frequency should not infers with usual communications devices. Energy efficiency is a critical point for the energetic life of an integrated and autonomous system. For this reason, many researchers work in order to find an optimal way to communicate between the device and the external world. There are three aspects of this research: the first one focuses on the silicon device technologies and materials. The second one focuses on the architecture trying to define the most efficient hardware architecture for

embedded computing integration. Finally, the third one focuses on communication protocol, as well at hardware level – antennas, computing, power consumption – at system level – soft radio, compression, computing complexity reduction.

The most important part of required computing capacity is devoted to video processing. It is crucial for a diagnostic helping device such an endocapsule. The practitioners have to visualize the body exploration which requires large computing capacity to ensure a confortable real-time high resolution video. Video processing is also essential for the control of mechanical parts of endocapsules that enable movement or biopsy. For example its purpose is to extract features from the image for positioning. Consequently, the video processing block usually requires large silicon area on the component. For this reason this chapter will focus in the video processing part.

2.2 Diagnostic helping devices: The needs

One of the domains that requires an efficient and accurate exploration is the endoscopy. The length of the digestive system causes many limitations. However they are drastically reduced from a decade thanks to the conception of the endoscopic video capsule.

Video endocapsules are shown to be useful in many cases:

- Unexplained digestive bleeding:

They represent about 5% of the digestive bleeding. general's methods offer very poor result. Profitability of radiological examination is only between 5 and 10% because no direct visualization of the mucous membrane is possible. Using enteroscopy, the profitability diagnosis for the lesions of hail is between 15 and 30% which is far from 100%. Using an endocapsule the profitability diagnosis are higher than that of the thorough enteroscopy, up to 70% (Maieron et al, 2004) (Fireman et al, 2004) (Selby et al, 2004).

- Crohn syndrom and hemorrhagic recto-colitis:

Chohn symdrom concerns about 2.5 billion patient in the world. This number increases each year. In this case, literature also shows that the endocapsule is a good choice for first intention exploration of clinical suspicion when traditional methods such as fibroscopy coloscopy and biopsies are negative (Bernardini, 2008).

- Polyps and hail tumors:

On 1042 examinations carried out, it was diagnosed 6-8% of tumors of hail (malignant 50%) (Lewis, Miami 2004). Endocapsule is especially efficient in the detection of small tumors (< 1 cm) which are difficult to see by general exploration such as simple radiological examination. There is also an interest like examination of tracking in the event of clinical suspicion (carcinoïde, lymphoma) due to the non-invasive nature of the technique and its simplicity of implementation for the patient.

By the literature, state of art, contact with practitioners and the study of diagnosis methods, diagnostic helping devices can benefit from following applications.

- Vision and real time 3D reconstruction of the scene is used to determine precisely the size of the lesions. This is used to find the optimal solution to treat the patient. At this time, the size of an anomaly is determined by the experience of the practitioner.

- Real time and autonomous detection of tumors, polyps, lesions and bleeding. Sometime, an anomaly can be very difficult to detect due to its localisation or its little size. The goal is to have the higher profitability diagnoses possible. This kind of processing is also very useful to determinate the region of interest - the region where an abnormally is seen - in the image. An autonomous detection should allow a better management of the power consumption by sending to the external world the image of the anomaly only.

- Spectrography is a possible solution to define the nature of a tumor when the biopsy not easily feasible. Spectrography is based on the spectral response of the organic fabric to a laser operating at a specific wavelength (Péry, 2008).

2.3 Algorithms used for general image processing in consumer's devices and diagnostic helping

The importance of the consumer devices market pushes the academic and industrial labs to innovate. This is required by the integration of brand new features in order to create new products, while maintaining the production cost as low as possible. Most of these new features require high computing capacities while silicon area must be kept under control and power consumption need to be sustained as low as possible. First, silicon area has a direct impact on production cost; moreover, too large components may be incompatible with a product form factor. Power consumption has a direct impact on battery life, which is crucial for handled products.

For example, on 2010, cell phones' image sensor represented about 80% of the overall sensor market for about 5 000 millions Dollar . These sensors are systematically associated with a digital Image and Signal Processor (ISP) to reconstruct and enhance the images from raw format. Cell phone integrators need video module, which include a video sensor and ISP at a price of about one dollar. Lenses and sensor costs are reduced as most as possible by reduction of the matrix and pixel size (today 2µm pixel are the state of the art). In addition to traditional color image reconstruction from raw data, this pixel size reduction implies an image quality degradation that must be corrected using digital ISP. An example of traditional image correction and reconstruction pipeline is presented in Figure 2.3. However to keep production costs low, their silicon area must be maintained under a few square millimetres using today's technologies. This forbids the use of traditional image processing approaches such as the use of a frame memory which may require more than times of silicon area budget.

Additional computing resources are used for high level application such as face recognition or augmented reality. Digital cameras and security cameras represent another part of the market of embedded image processing. Depending on their usage, they can embed low level to complex high level algorithms, from simple image enhancement to face recognition or motion detection and tracking.
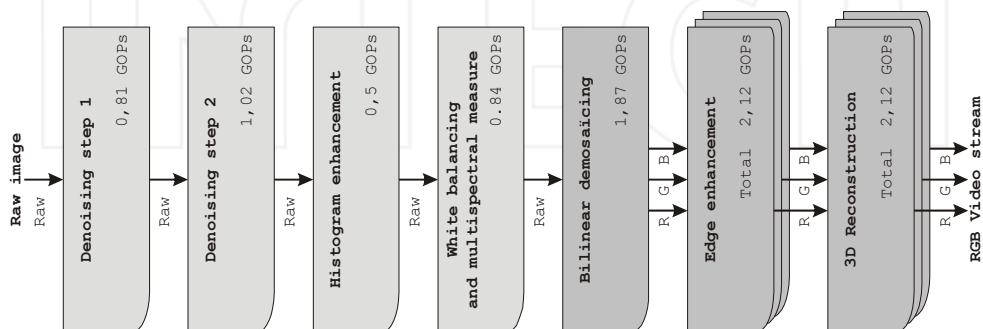

Fig. 2.3. Example a of a low level image reconstruction video pipe.

Todays handled video games and digital cameras are able to handle 3D as well for image grabbing and displaying. Designers now consider this feature must be integrated into devices. This feature requires specific algorithms to process images, especially when they are grabbed by a stereoscopic pair.

Basic image enhancement algorithms are used as well for image grabbing as for image displaying. Depending on the nature of the targeted application, high level algorithms may be used in addition. For example, interest point detection is widely used for face detection or augmented reality. Stereoscopy may be also used for this last purpose.

2.4 Algorithms used for diagnosis helping

By the analysis of the applications needed to enhance the diagnosis, it is possible to define a selection of video-processing algorithms. If we let on the side the most common processing used for image reconstruction and enhancement, which is the first step of all image acquisition, one can extract the following algorithms:

- Shape detector:

In order to define a region of interest in the image, this kind of detector is very common. In a simplifying way, we can summarize this algorithm by the analysis of the reflectance or depth discontinuity in an image; actually, the intensity discontinuity allows the edge definition. The principle of edge detection is based on the study of the derivative of the intensity function in the image: local extrema of the gradient and passages by zero of the Laplacian. This can normally be achieved by convolution like approaches;

- Colour analysis:

This technique allows to fetch the information about the incident spectrum wave. This is similar to spectroscopy. The base of this method is to record a certain color profile and to compare it to a matching table, which contains the known color profile. This enables to find the needed information for example the nature of a tumor – considering that each tumor has a specific color response.

- Labelling:

The goal is to give an identification code to each region of interest in the image in order to process them separately; the labelling is one of the most important processing with the form recognition. If we simplify to the maximum, this method is based on the scanning of the image, each time that a region of interest is found, which was defined by a previous processing like form recognition, a label can be attributed. There are many different technique of labelling, depending of the complexity of the image. Graba (Graba, 2006) proposes a solution to integrate labelling in a small 3D vision sensor. Lacassagne (Lacassagne, 2009) proposes an extremely fast method to process the labelling.

- 3D reconstruction:

The depth reconstruction is usually based on three different solutions: the so called active one, based pattern projection read by a camera. The second one is passive stereoscopy, with two or more cameras allowing a triangulation from the images (Darouich, 2010). N. Ventroux and R. Schimit (Ventroux, 2009) defines a solution to achieve a 3D reconstruction device based on stereoscopic method for autonomous cars. Kolar (Kolar, 2007) (Kolar, 2009) defines a way to integrate the 3D reconstruction into an integrated vision sensor for an endoscopic video capsule. Ruben Machucho-Cadena and Eduardo Bayro-Corrochano (Machucho-Cadena, 2010) present a solution to create a 3D model of a brain tumor from endoscopic and ultra-sound images. The processing will depend on the complexity of the scene and the required precision. The third solution is based on the time of fly of an energetic wave (Oggier, 2004).

- Form recognition and classification:

The form recognition allows finding a certain object from the raw data in order to classify it and to take a decision; this can be to stop your camera-equipped car when an obstacle is detected (Ponsa and al., 2005). This method is based on two different steps: firstly, the system needs to learn what kind of object it has to detect. This is usually done by a method called AdaBoost. A database that contains the objects to recognize is used to define classifier coefficients in order to obtain the good set of output. Finally, a classifier is able to determinate what kind of object is present on the raw data and to classify it.

The analysis of diagnostic helping algorithms shows that simple algorithms such as pattern recognition or stereo reconstruction are required. These approaches require a computing capacity of hundred billion of operations (GOPs) in order to be executed. Moreover, some of them also may require a frame memory to be correctly executed. Figure 2.4 shows an approximation of computing capacity required expressed in GOPs of the previously presented algorithms.
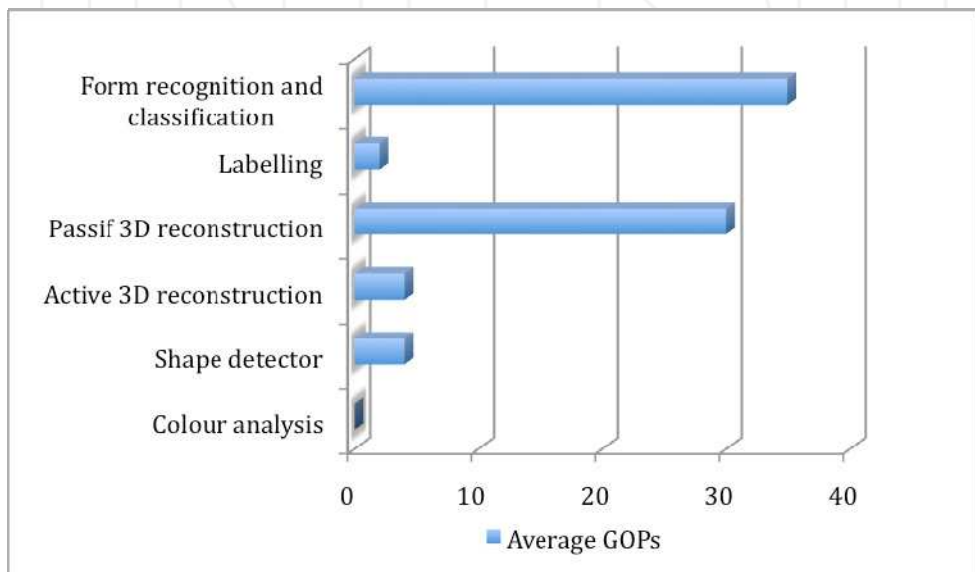

Fig. 2.4. GOPs consumption of diagnosis helping algorithms

But one of the most interesting thing that we can see after the analysis of the algorithms used for the diagnosis helping, is that we can find the same algorithms in consumers devices likes smartphone, camera, game station, etc. Innovations concern architecture design as well as algorithmic definition. Researches also involve co-design and high level synthesis in order to match embedded systems constrains. The expertise of image processing community and embedded devices is widely used for consumers devices researches.

3. Application to endoscopic imaging

A co-design approach is required in order to meet the computing resources requirement of handled diagnostic devices. These approaches are widely used in the community of

consumer's devices. First, the whole application set is studied in order to define computing intensive blocks from image processing applications. The first section presents them and their operators that can be ported to hardware resources for both medical and consumer's application. The second section presents a brief state of the art of hardware components known to be efficient for embedded computing intensive image processing from both industrial and academic works. Finally, third section presents a feasibility study of an autonomous endoscopic capsule which has not only the ability to grab and outcast videos like today's one, but also to process them in order to emphasize specific medical abnormality.

3.1 Required operations for image processing

Study of the applications done in previous part of this chapter gives a set of atomic operators. The first processing level needed for every sensed picture consists in a low level image reconstruction and enhancement as previously presented in Figure 2.3. It is realized by algorithms that are pipelined downstream of the image sensor.

In order to capture a correct image, exposure metering and system for auto-focusing must take place. The second step is devoted to the elimination of the electronic noise, which degrades the signal. A contrast enhancement step permits a better usage the sensor dynamic range. Because many types of illuminant sources induce color variation, white balancing makes image colors look natural. The demosaicing step interpolates a complete color image from raw data produced by a color-filtered sensor such as Bayer filter. Finally, various image enhancement processes, such as distortion correction or adaptive edge and contrast enhancement can be applied. The last step (not discussed in this paper) is devoted to the compression and the storage of the image, or to detect points of interest such as corners facilitating object recognition.

- Image capture

Fine exposure-metering methods are required to ensure a correct use of the sensor dynamic range. Similar methods can be employed to ensure that the subject is correctly focused and suitably sharp.

- Exposure control

This step consists in defining exposure parameters which are exposure time, optical aperture, linear ISO sensor sensitivity, and scene luminance.

In smart phones, but also in non Single Lens Reflex (SLR) cameras and camcorders, exposure control can be achieved by direct analysis of the stream of pictures from the sensor as done by Shimizu et al. (Shimizu, 1992).

- Auto-focusing

Auto-focusing consists in measuring image sharpness in a region of interest while displacing certain optical elements. The most common methods are either gradient-based or Laplacian-based such as (Lee, 1980). The region of interest is conventionally considered to be the centre of the image.

- Noise reduction

The use of multiple mega-pixel sensors is encouraged by the current market trends for mobile devices. This tendency has also led to reduction in pixel size, thereby limiting both SNR and overall image quality as explained in Chen et al. (Chen, 2000). Some of the correctible noise is especially due to the CMOS technologies used in image sensors.

- Pixel noise is directly correlated with photo site area, since photodiode voltage following exposure must be comparable to voltage value after reset (if the latter is more than zero, reset is incomplete). The resulting noise, which can be significant,

takes the form of a residual current generated when a pixel is read quickly. Pixel noise is also caused by thermal excitation and leakage. Spatial and temporal disparities caused by such noise are observable and can be statistically characterized.

- Amplification and quantization noise is directly due to ADC sampling. In CMOS sensors, an amplifier and an ADC are present for each column. As in any other electronic device, the signal generated by them includes thermal noise to which quantization noise must also be added.

It is possible to reduce the impact of amplification and quantization noise on images in various ways. The first is to cancel Fixed Pattern Noise (FPN) by deleting characterized noise pixel-per-pixel or column-per-column. The second is to replace any absurd pixel values, which are also those most visible to the human eye. This can be done using Gaussian–kernel convolution or adaptive filtering, for example with bilateral filters (Tomasi, 1998).

- Contrast enhancement

This step allows an optimum use of the full dynamic range of the image. Histogram equalization can be applied to the whole image. The existing literature also describes various embeddable, local adaptive methods. These methods, like High Dynamic Range Imaging (HDRi), are used to extract high and low light values that are not visible on standard displays. Numerous signals are recorded by the sensors in dark and bright areas of the image. Without tone mapping, these signals are not visible on a standard monitor due to saturation effect. Adaptive methods ensure local contrast enhancement using local gamma, local histogram or Retinex-like approaches.

- White balancing and multispectral analysis

Sensor pixels are covered by a color filter such as the well known Bayer one that they ''grab'' signals corresponding to each primary color. This allows measurement of the absolute luminance values for each color component. These values depend on the scene illuminant color, which induces a global image color—yellow-orange for tungsten and blue-violet for fluorescent light sources. This step aims to determine illuminant color and obtain realistic image colors. The best known method is the grey world assumption, which is used in numerous applications and may vary to other methods like the grey-edge one as proposed by van de Weijer and Gevers (Weijer, 2007).

Multispectral analysis consists in lighting the scene using different wavelength. Nature of the object may be determined by analysis of its response to these different lights. For example a some kind of tumour would be revealed by a 1200 to 1400 nm wavelength.

- Color plan interpolation

The crucial demosaicing step computes each RGB or YUV plan from a single raw image ''grabbed'' by the sensor, like any camera. There is literature available on a large number of research projects relating to this step, such as. While simple bilinear interpolation calls for computing pixel values by averaging the neighbourhood, other methods use channel-to-channel correlations or edge-of-neighbourhood to adapt the demosaicing method to neighbourhood content.

- Image enhancement

Enhancement is necessary to ensure a high quality image. A good contrast balance and sharp edges are two essential parameters for visual perception of an image. Therefore, they can be corrected at the same time. Although correct exposure allows efficient use of the sensor dynamic range, histogram-based processing, like normalization and equalization, are also used to enhance dynamic range. Such processing usually takes place after noise reduction. Edge enhancement can then be performed with a high-pass filter. For this

purpose, convolution-based filters like the Sobel filter, unsharp mask or Canny Deriche can be used, as can local adaptive filters, which serve to sharpen images. Image enhancement is traditionally executed in spatial domain, but new approaches tends to execute process in wavelet domains (Courroux, 2010).

- Pattern Recognition

Any device that need to detect specific feature in an image such as face recognition and smile detection like most digital cameras must detect interest points or shape (red eye, face, smile). Traditional methods can be used, however, new approaches based on dynamic neural network are under study (Bichler, 2011).

- Tracking

Many consumer devices are able to detect and to track moving objects such as faces. This is the case for video-conferences devices or digital cameras that uses this feature to enhance auto-focusing. Methods that allow object tracking can be based on feature detection. For example the Harris (Harris, 1998) corner detector. This algorithm is based on three convolutions that process horizontally and vertically edge filtering. The detection of the corner is allowed by the overlapping of the previous results. A final step consists in a cleaning filter to keep only the righteous interest points.

Global exposure control	< 1 MOPs
Autofocus (spot)	< 1 MOPs
FPN removal	0.210 GOPs
White balancing and multispectral detection	20 MOPs
Convolution 3×3	2.5 GOPs
Demosaicing	1.2 to 3 GOPs
Image enhancement	3 GOPs
Active 3D reconstruction	1.5 GOPs
Labelling	2 GOPs + frame memory
Object recognition (tumor, polyp etc)	4 to 30 GOPs + frame memory
TOTAL	~40 GOPs

Table 2.1. Example of the required computing capacity for low level image processing.

Previous approaches have presented image and signal processing algorithmic. They can be ported onto programmable or configurable components on the shelves, Application Specific Processors (ASIPs) may be designed for the execution of the algorithms, or they can be hardwired. The choice of the hardware implementation depends on the constraints to meet for the targeted design. Table 2.1 shows an example of different computing resources that are required to process some of most common low level image processing. It shows the variety of approaches and the variety of required resources.

3.2 A brief survey of embedded computing architectures

Consumers devices such as smart phone, cameras and handled devices drive a large market. This is especially the case for embedded real-time video processing that is the subject of both academic and industrial researches. These researches are driven by the market constraints. First the silicon area infers the component cost, next the power consumption determines if this component can cope with battery powered devices. Finally, flexibility is a feature that is more and more required by integrators. This allows them to use the same component in

different generation of devices by simply reconfiguring the hardware or by an update of the firmware or the software of the devices' components. As the choice of an hardware implementation for signal processing can be complex depending on the silicon area constraints, power consumption and computing capacity requirement of the applications. This section presents some of the architectures that may enable image enhancement on smart phone, considering their complexity in terms of gates count or silicon area, their power consumption and their ability to run different kind of processing. Many classifications of these signal processing architectures can be done. For didactic purposes, this section split them into three parts. Dedicated architectures are firstly presented, followed by reconfigurable architectures and by programmable architecture.

A. Dedicated architectures

Are considered as dedicated architecture, components that are made of specialized wired operators grouped together in order to realize more complex hardwired functionalities. These architectures are low silicon footprints and are usually low-power, thus enabling them to be used inside embedded systems such as cell phones. Indeed, their fully wired design is optimized for the applications integration constraints. Today, they are often used by integrators for low level pixel processing such as contrast and color correction, demosacing (Garcia-Lammond, 2008) or denoising (P.Y. Chen, 2008). Designers group these Intellectual Properties (IPs) to forms complete signal processing architecture such as a video pipe image enhancement. For example (Zhou 2003) architecture is able to process Video Gate Array (640×480 pixels (VGA) video stream at 30 frames per second (fps), while Hitachi (Nakano 1998) proposes a component that is able to process Super eXtended Gate Array (SXGA) pictures. However, these more complex systems require an external memory acting as a frame buffer to work properly. Videantis proposes two processors (Videantis inc., 2007) (Videantis inc., 2008) that are able to process High Definition (HD) video stream conforming to standards HD 720p and HD 1080p. The most powerful of them requires large silicon area and power consumption which is not compatible with their integration into low-cost components. As dedicated operators cannot be autonomous, they need to be used in association with embedded processors (e.g. ARMs or MIPSs) and an external memory or finite state machines. This is a common solution for low-power mobile devices like cell phones or compact cameras. Despite the high computational efficiency of these solutions, they lack flexibility due to their hardwired implementation that allows to the customers to configure only a set of limited predefined parameters, these solutions are widely used thanks to a short time to market.

B. Reconfigurable architectures

Reconfigurable architectures may be seen as evolutions of dedicated operators, especially when they are used in complex System-on-Chips (SoCs). SoCs need of flexibility and operator reuse for different applications pushes the architect to define methods for this purpose. For example, the Coarse Grained Reconfigurable Image Processor (CRISP) architecture (Chen, 2008a) can handle HD 1080p video streams. It was specifically designed in order to run image processing and enhancement application downstream the image sensor with more flexibility than dedicated IPs. However supported processes are limited by hardwired modules that compose the design. It also was designed to limit its silicon area usage and power consumption in order to be embeddable into smart phones. Its implementation requires approximately 170 kGates and 74 kb of memory. This

correspond to a 400 kGates and 5 mm2 when implemented in 180 nm technology – an extrapolation gives about 1 mm2 of silicon area in Taiwan SeMi Conductor (TSMC) 65 nm. Its given power consumption is 218 mW at 115 MHz while it can run a complete image processing on HD 1080p video streams at 55 fps. Unfortunately, its flexibility is limited by its hard-wired embedded processes. Moreover, to run algorithms properly, it must be associated with memory resources. DART (David, 2002), MORA, MorphoSys or ADRES approaches can be cited, however, more flexible reconfigurable architectures are, and more fine grained their reconfigurability is. The reconfigurability elements of such architecture, especially interconnects, implies an important silicon area overcost, thus can be larger than the computing elements themselves making their integration into low-cost devices difficult.

C. Programmable architectures

Programmable architectures can be seen as specifically designed fine grained reconfigurable architectures. In order to maintain a low silicon area and high computing performance over power consumption, architects have to specialize their design for an application predefined set. Spiral Gateway, for example, proposes RICA, a configurable System on Chip (SoC), which is based on algorithm analysis (Khawam, 2008) and is thus programmable within the scope of the initial application set. Tensilica provides another product that is extended instruction set processors (Tensilica). SiliconHive markets a processor template that is customized by application code analysis. Its type and number of operators – from 4 to 128 – can be customized at the time of chip design. For the automotive market, NEC has devised the ImapCar processor (Kyo, 2005) containing 128 SIMD – Single Instruction Multiple Data means that every processor executes the same instruction on different data, for example each processor do the same job on each pixel of an image – parallel arithmetic and logic units with a power consumption of more than one Watt. Xetal also proposes a programmable, massively-parallel processor integrating 320 computing units (Abbo, 2008). SIMPil (Gentile, 2005) architecture calls for parallelized 4096 processors, each of which is intended to compute a single pixel block. Stream Processors Inc., a commercial spinoff of Stanford's Imagine project (Stream, 2007) and Massasshusset Institute of Technology (MIT), proposes STORM, a family of parallel chips that can handle video streams. These components are not directly embeddable in cell phones due to their high power consumption and large area. An acceptable silicon "budget" is about 1 to 2 mm2 in a typical 65 nm technology with a power consumption of less than half a watt. These constraints is lacking for programmable architectures in this competitive market niche.

However, the common feature in all these programmable components is the use of different forms of parallelism such Single Instruction Multiple Data (SIMD) and Very Long Instruction Word (VLIW), making them efficient for computing regular data patterns. This is especially the case for stream processors. This brief study of the state of the art architecture shows that many of the most efficient flexible machines are based on multiple programmable processors running in SIMD mode. Moreover, VLIW processors are often used allowing the ILP of programs to be exploited. In this fact, the proposed architecture includes these features (programmability, SIMD and VLIW). However, data access remains an important bottleneck that limits computing bandwidth. In order to get a high computing capacity, the proposed architecture is designed to separate data access and computing, in this way, we can achieve the computation directly on incoming video stream without needing an external frame buffer.

3.3 Proposed vision architecture for integrated diagnostic helping devices

The proposed architecture is based on the eISP (Thevenin, 2010) processor that is designed for smart phone embedded video and is derived to give enough computing capacity to support diagnostic helping image processing algorithms that could be required in an endocapsule. Our study established an approximation of the required computing capacity of about 50 GOPs for an average power consumption of less than a half Watt, and a maximum silicon area of 15 mm² dedicated to computations.

As shown previously, algorithms can easily be divided in elementary stages and pipelined. One of the most efficient architecture models consists in splitting a whole multiprocessor architecture into elementary computing tiles as shown in Figure 3.1. Each of them acts as an autonomous SIMD computer that can execute a process. Figure 3.2 depict a P processors computing tile. Each computing tiles is connected using a bus, allowing the execution of different kind of processes. For example, video processing are chained as shown in the first section can be mapped onto each computing tile.
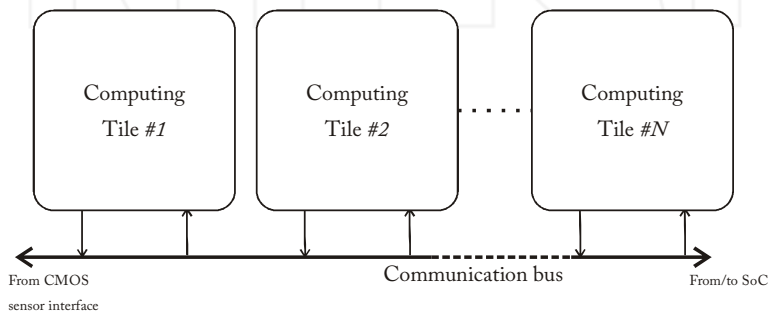

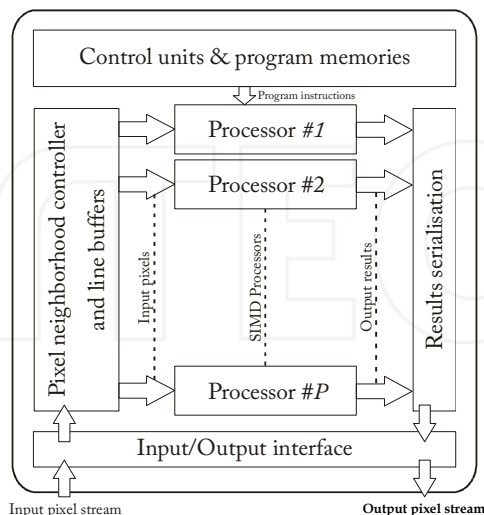Fig. 3.1. eISP, a compultile tile architecture.

Fig. 3.2. A P processors computing tile.

Different instances of computing tiles are characterized in terms of computing capacity, power consumption, silicon area in function of their number of processor and memory resources. An example characterization of the architecture is shown on Figure 3.3. This work gives a normalized performance measure expressed in MOPs/mW and GOPs/mm². Standard instance of the eISP architecture gives a computing capacity of about 25GOPs/mm² for 100mW. Reaching a computing capacity of 100 GOPs that would be required for image processing in diagnostic helping device would require 4mm² of silicon area and 400mW of power consumption.

Each computing tile can be generated with a set of parameters that are given by the designer. For example the data-path width, usually 8 to 32 bits and its operators, memory maps, that is distributed in each processor or that is shared with all processor of a same computing tile.

Sizing the whole architecture depends on the total required computing capacity, but also on the computing capacity that the designer need for each task that will be ported on each computing tile. Designer may uses results of the characterization, as the example shown on Figure 3.3 to size its architecture. He can generate computing tiles and connect them to the communication bus. Final synthesizes and simulations are required to check the designed architecture. Finally, the eISP can be integrated into a complete System on Chip or to a Lab on Chip that include control and communication components.
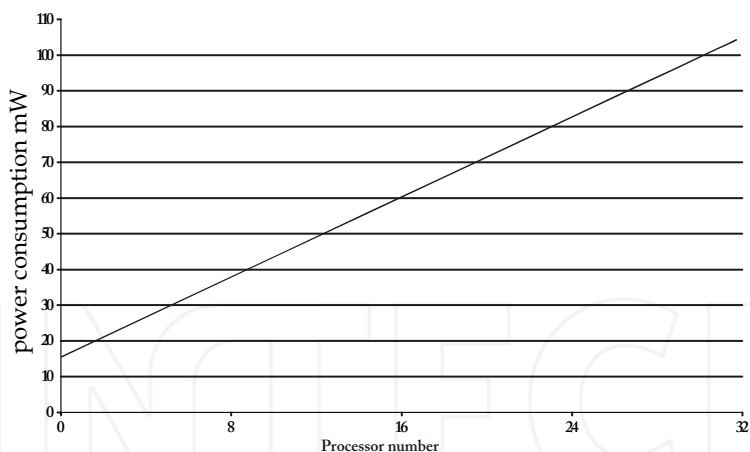

Fig. 3.3. Characterization of the power consumption of a single computing tile eISP architecture versus number of processors.

A complete characterization of the eISP architecture in TSMC 65nm was done allowing an accurate design space exploration. We can add up to two frame buffer for HD 720p require that would require 4 mm² for each frame. Thus, allow high level processing such as video compression and labelling that requires up to several dozen GOPs and a frame memory depending on the selected implementation.

4. Conclusion

This chapter has presented the algorithms that could be used for digital image processing in handled diagnostic devices, and more precisely in the case of endoscopy. As research in consumer devices imaging is intense, a comparisons of the algorithms that are used in that domain is done in this chapter. This work shows similarities between the approaches. These similarities can be exploited in order to transfer the hardware processors initially designed for consumers market – such as cell phone or gaming – to integrated medical domain. The case of the endoscopic video capsule is used due to its highly constrained integrability, as well in terms of silicon area or power consumption and computational capacity. A state of the art of the architectures that could match these constraints is described. It shows that the existent architectures do not to perfectly cope with computational requirement, silicon area or power consumption. A computing architecture derived from the eISP, an image signal processor designed for low level image enhancement is proposed. With less than 5 mm² and 0.5 Watt of power consumption, this can integrate the required computing and memory resources for handled diagnostic device in limited constraints inherent to this domain. Due to its programmability, it can be used not only as image enhancement architecture, but also as a high-level diagnostic helping processor by executing processes like form recognition, 3D-reconstruction, shape detector etc.

The use of such signal processing architecture in conjunction with complete robotized diagnostic helping platforms as (Valdastri, 2009) may allows the conception of an autonomous lab-on-chip that would be able to execute simple tasks like free move and biopsy.
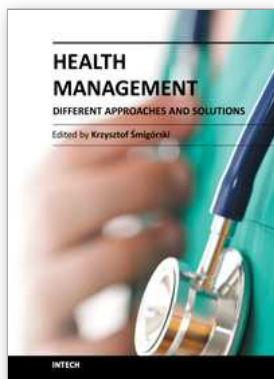
5. References

A.A. Abbo, R.P. Kleihorst, V. Choudhary, L. Sevat, P.Wielage, S.Mouy, B. Vermeulen and M. Heijligers. (2008) Xetal-II: A 107 GOPS, 600 mW Massively Parallel Processor for Video Scene Analysis. Solid-State Circuits, *IEEE Journal of, vol. 43, no. 1*, pp 192–201, Jan. 2008.

F. Bernardini, M. Cerbo, T. Jefferson, A. Lo Scalzo, M. Ratti, (2008). Age.na.s HTA Report - Wireless Capsule Endoscopy in the diagnosis of small bowel disease, Rome, September 2008.

O. Bichler, D. Querlioz, S.J. Thorpe, J.P Bourgoin, C. Gamrat. (2011) A wavelet-based demosaicing algorithm for embedded applications; *International Joint Conference on Neural Networks (IJCNN - 2011)*, San José, Etats-unis, 31/07/2011 - 05/08/2011.

T. Chen, Peter Catrysse, Abbas E Gamal and Brian W. (2008) How small should pixel size be ? In *Proceedings of SPIE, April 2000, vol 7, no. 9*, pp 451–459, 2000.

J.C. Chen and Shao-Yi Chien (2008). CRISP: Coarse-Grained Reconfigurable Image Stream Processor for Digital Still Cameras and Camcorders. *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp 1223–1236, Sept 2008.

P.Y. Chen, Chih-Yuan Lien and Yi-Ming Lin. (2008) A real-time image denoising chip. In Circuits and Systems, *ISCAS 2008. IEEE International Symposium on*, pp 3390–3393, May 2008.

S. Courroux, S. Guyetant, S. Chevobbe S., M. Paindavoine. (2010), Reconfigurable Computing: Architectures, Tools and Applications, *International Conference on*

Design and Architectures for Signal and Image Processing (DASIP – 2010), Edimbourg ; Royaume-uni, 2010.

M. Darouich, S. Guyetant and D. Lavenier. (2010) A Reconfigurable Disparity Engine for Stereovision in Advanced Driver Assistance Systems *Lecture Notes in Computer Science*, , Volume 5992, 2010.

R. David, D. Chillet, S. Pillement, O. Sentieys. (2002) DART: a dynamically reconfigurable architecture dealing with future mobile telecommunications constraints, *Proceedings International Parallel and Distributed Processing Symposium, IPDPS 2002*, pp. 156, 2002.

Fireman and al (2004), *Eur J GEH* 2004

J. Garcia-Lamont, M. Aleman-Arce and J. Waissman-Vilanova. (2008) A Digital Real Time Image Demosaicking Implementation for High Definition Video Cameras. *In Electronics, Robotics and Automotive Mechanics Conference, 2008. CERMA '08*, pp 565–569, 30 2008-Oct.

A. Gentile, S. Vitabile, L. Verdoscia and F. Sorbello. (2005) Image processing chain for digital still cameras based on the SIMPil architecture. *Parallel Processing, 2005. ICPP 2005 Workshops. International Conference Workshops on,* pp 215–222, June 2005.

P. Gomes, (2011) Surgical robotics: Reviewing the past, analysing the present, imagining the future, *Robot. Comput.-Integr. Manuf.*, vol. 27, no. 2, pp. 261-266, Apr. 2011.

T. Graba (2009), Etude d'une architecture de traitement pour un capteur intégré de vision 3D, *Phdthesis, Université Pierre and Marie Curie*, 2009.

K. Harada, E. Susilo, N. Ng Pak, A. Menciassi, and P. Dario, (2008) Design of a Bending Module for Assembling Reconfigurable Endoluminal Surgical System Pisa, *ISG conference,* Tuscany, Italy - June 4-6, 2008.

Harris and Stephans, (1988) A Combined Corner and Edge Detector. *In Alvey Vision Conference*, pp 147–152, 1988.

M. Hartmann, V. Pantazis, T. Vander Aa, M. Berekovic, C. Hochberger and B. de Sutter, (2007) Still Image Processing on Coarse-Grained Reconfigurable Array Architectures. *In Embedded Systems for Real-Time Multimedia, 2007. ESTIMedia 2007. IEEE/ACM/IFIP Workshop on*, pp 67–72, Oct. 2007.

R. Machucho-Cadena and E. Bayro-Corrochano, (2010) 3D Reconstruction of Brain Tumors from Endoscopic and Ultrasound Images, *Pattern Recognition Recent Advances*, *InTech,* Adam Herout (Ed.), ISBN: 978-953-7619-90-9, , 2010.

A. Menciassi, C. Stefanini, G. Orlandi, M. Quirini, P. Dario, (2006) Towards active capsular endoscopy: preliminary results on a legged platform, *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS '06*, Page(s): 2215 – 2218, 2006.

M. Katona, A. Pižurica, N. Teslic, V. Kovacevic and W. Philips. (2006) A real-time wavelet-domain video denoising implementation in FPGA. *EURASIP J. Embedded Syst., vol. 2006, no. 1*, pages 6–6, 2006.

A. Karargyris, N. Bourbakis, (2010) Wireless Capsule Endoscopy and Endoscopic Imaging: A Survey on Various Methodologies Presented, *IEEE Engineering in Medicine and Biology Magazine, Vol. 29 Issue:1* ,pages 72 - 83 , Jan.-Feb. 2010

S. Khawam, I. Nousias, M. Milward, Ying Yi, M. Muir and T. Arslan. (2008) The Reconfigurable Instruction Cell Array. Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, vol. 16, no. 1, pages 75–85, Jan. 2008.

T. S. Kim, S. Y. Song, H. Jung, J. Kim and E.-S. Yoon (2007), Micro Capsule Endoscope for Gastro Intestinal Trac, *IEEE EMBS,* 2007, pp 2823-2826.

A. Kolar, O. Romain , T. Graba, T. Ea and B. Granado, (2008) The Integrated Active Stereoscopic Vision Theory, *Integration and Application Stereo Vision, InTech*, ISBN 978-953-7619-22-0, November 2008

A. Kolar, A. Pinna, O. Romain, S. Viateur, T. Ea, E. Belhaire, T. Graba and B. Granado (2009), A multi shutter time sensor for multi-spectral imaging in a 3D Reconstruction integrated sensor, *IEEE Sensor Journal,* vol 9, pp 478-484, 2009

S. Kyo, S. Okazaki and T. Arai, (2005) An integrated memory array processor architecture for embedded image recognition systems. *Computer Architecture, 2005. ISCA '05. Proceedings. 32nd International Symposium on,* pp 134–145, June 2005.

J.S. Lee. (1980) Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence., vol. PAMI-2*, pp 165–168, March 1980.

L. Lacassagne, B. Zavidovique, (2009) Light Speed Labeling for RISC architectures, *16th IEEE International Conference on Image Processing (ICIP),* 2009 , Page(s): 3245 - 3248

Ming-Hau Lee, Hartej Singh, Guangming Lu, Nader Bagherzadeh, Fadi J. Kurdahi, Fadi and J. Kurdahi. (2000) Design and Implementation of the MorphoSys Reconfigurable Computing Processor. *In Journal of VLSI and Signal Processing-Systems for Signal, Image and Video Technology.* Kluwer Academic Publishers, 2000.

Maieron and al, (2004) *Endoscopy* 2004

N. Nakano, R. Nishimura, H. Sai, A. Nishizawa and H. Komatsu, (1998) Digital still camera system for megapixel CCD. *Consumer Electronics, IEEE Transactions on, vol. 44, no. 3*, pages 581–586, Aug 1998.

T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, Lustenberger, F. & Blanc, N., (2004) An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger),*SPIE, Optical Design and Engineering,* pp534-545, 2004.

E. Péry, (2008) Spectroscopie bimodale en diffusion élastique et autofluorescence résolue spatialement: instrumentation, modélisation des interactions lumiére-tissus et application à la caractérisation de tissus biologiques ex vivo et in vivo pour la détection de cancers*, Phdthesis, Institut National Polytechnique de Lorraine*, 2008.

D. Ponsa, A. L´opez, F. Lumbreras, J. Serrat, T. Graf, (2005) 3D Vehicle Sensor based on Monocular Vision, *IEEE Conference on Intelligent Transportation Systems*, 2005.

M. Quirini, S. Scapellato, P. Valdastri, A. Menciassi and P. Dario, (2007), An Approach to Capsular Endoscopywith Active Motion *IEEE EMBS,* 2007, pp 2827-2830.

Selby and al, Gastrointest Endosc 2004.

Tensilica Co. (2007) 388VDO Video *DSP Product Brief.*, Tensilica Co., 2007.

S. Shimizu, T. Kondo, T. Kohashi, M. Tsurata, T. Komuro, (1992), A new algorithm for exposure based on fuzzy logic for video cameras , *IEEE Transactions on Consumer Electronics Volume: 38 , Issue: 3*, 1992 , Page(s): 617 – 623.

Stream Processors, Inc. (2007) Storm-1 Stream Processors, SP16HP-G220 *Product Brief. Stream Processors, Inc.,* Apr. 2007.

M. Thevenin, M. Paindavoine, L. Letellier, R. Schmit, and B. Heyrman, (2010) The eISP a low-power and tiny silicon footprint programmable video architecture, *Journal of Real-Time Image Processing,* pp. 1–14, Jun. 2010.

Texas Instruments (2006) Texas Instrument TMS320DSC21 : A High-Performance, Programmable, Single Chip Digital Signal Processing Solution to Digital Still Cameras.

C. Tomasi and R. Manduchi, (1998) Bilateral Filtering for Gray and Color Images, Proceedings of the *1998 IEEE International Conference on Computer Vision*, Bombay, India, 1998.

P. Valdastri, R. J. Webster III, C. Quaglia, M. Quirini, A. Menciassi, and P. Dario (2009) A New Mechanism for Meso-Scale Legged Locomotion in Compliant Tubular Environments. *IEEE Transactions on Robotics,* 2009.

Videantis Inc. (2007) v-MP2000SD, Dual-Core Multi-Standard Video Codec IP Solution. *Technical Report, Videantis Inc.,* 2007.

Videantis Inc. (2008) v-MP4180HDX, Full HD 1080p Video Codec Integrated Solution. *Technical Report, Videantis Inc.,* 2008.

J. van de Weijer, A. Gijsenij and Th. Gevers. (2007) Edge-based color constancy. *IEEE Transactions on Image Processing,* 2007.

Rongzheng Zhou, Xuefeng Chen, Feng Liu, Jie He, Tiankang Liao, Yanfeng Su, Jinghua Ye, Yajie Qin, Xiaofeng Yi and Zhiliang Hong. (2003) System-on-chip for mega-pixel digital camera processor with auto control functions. *In ASIC, 2003. Proceedings. 5th International Conference on, volume 2,* pp 894–897 Vol.2, Oct. 2003.

Rongzheng Zhou, Xuefeng Chen, Feng Liu, Jie He, Tiankang Liao, Yanfeng Su, Jinghua Ye, Yajie Qin, Xiaofeng Yi and Zhiliang Hong. (2003) System-on-chip for mega-pixel digital camera processor with auto control functions. *In ASIC, 2003. Proceedings. 5th International Conference on, volume 2,* pp 894–897 Vol.2, Oct. 2003.

N. Ventroux, R. Schmit, F. Pasquet, P.-E. Viel, S. Guyetant. (2009) Stereovision-based 3D detection for automotive safety driving assistance, *12th International IEEE Conference on Intelligent Transportation Systems, 2009. ITSC '09,* pp 1, 4-7 Oct. 2009.

Health Management - Different Approaches and Solutions

Edited by Dr. Krzysztof Smigorski

The development in our understanding of health management ensures unprecedented possibilities in terms of explaining the causes of diseases and effective treatment. However, increased capabilities create new issues. Both, researchers and clinicians, as well as managers of healthcare units face new challenges: increasing validity and reliability of clinical trials, effectively distributing medical products, managing hospitals and clinics flexibly, and managing treatment processes efficiently. The aim of this book is to present issues relating to health management in a way that would be satisfying for academicians and practitioners. It is designed to be a forum for the experts in the thematic area to exchange viewpoints, and to present health management's state-of-art as a scientific and professional domain.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds