

# On the Feasibility of Miniaturised Vision Systems

Stephanie McBader ([mcbader@neuricam.com](mailto:mcbader@neuricam.com))

NeuriCam S.p.A, Via S M Maddalena 12, 38100 Trento, Italy

## Abstract

Compact vision systems capable of capturing images at high speed and interpreting them to give users high-level information on the scene are increasingly becoming available to serve a range of real-time applications. This paper discusses the reality of miniaturised vision systems-on-chip, the challenges which face their design, and their feasibility using modern and perhaps future technologies.

## 1 INTRODUCTION

Smart cameras are implementations of distributed computer architectures based on the principle that digital data processing is less costly than data transmission and therefore information should be converted to digital format and processed digitally as close as possible to the sensor. As technology progresses, the compactness of vision systems increases. Several independent advances in microelectronics, packaging and optics have contributed to the possibility of implementing complete digital vision systems on a single chip [1]. These include the very fast evolution of microelectronic components, brought about by better silicon fabrication processes, advances in digital design techniques, mixed analogue and digital design on the same silicon microchip, and the integration of optical sensors with other circuit elements in the widely available CMOS technology [2]. Powerful processing architectures can be embedded monolithically with the imager at a relatively low marginal cost [3]. Computer vision techniques have been improved for a number of years to yield efficient filtering, recognition and classification algorithms, while electrical interfaces and protocols have been defined for the convenient exchange of data.

This paper discusses the feasibility and challenges facing the design of single-chip vision systems, using examples of existing architectures and future proposals.

## 2 THE PAST

The consumer market sees many solutions to building real-time, pervasive vision systems. Existing vision systems come in various sizes and processing powers. While some occupy very large areas, such as an entire room or hall [4], others are simply a collection of modules in a single casing [5]. The requirement for the system's size, speed and power consumption is often application dependent; however, the idea of complete system

integration on a single chip has long been sought by researchers and developers [6]. To this extent, many processing architectures, platforms and cores have been proposed, developed and implemented with the prospect of incorporating them onto future single-chip solutions [7,8].

## 3 THE PRESENT

The current state-of-the-art in vision systems design follows two main advances in the electronics industry: That of embedded systems and mixed-signal designs, and that of improved CMOS pixel technology. Their merging gives birth to the notion of Vision Systems-on-Chip, which, by combining sensing and processing circuits on the same substrate, reduce size, power consumption, communication delays and component cost. An example of such miniaturised vision systems is the VISoc chip [9], developed at NeuriCam S.p.A, in Italy. The single-chip system integrates an optical sensor with pixel selection and signal conditioning circuits, a RISC processor, a neural processor, memory management logic and suitable interfaces.

Sensor (Vista)	320x256 APS
Frame Rate	180 FPS
Resolution	10-bits
On-Chip Processing	32-bit RISC
On-Chip Memory	93.5 Kbits
Off-Chip Memory	512KB FLASH 1MB SRAM
Neural Network	32-node
Clock Frequency	≤ 60 MHz
Processing Power	60 RISC MIPS 300 MACS
Power Dissipation	≤ 1 W
Process	0.35 μm
Size	< 36 mm <sup>2</sup>

Table 1: VISoc Characteristics

## 4 THE FUTURE

The VISoc chip is only one example illustrating the merits of modern day CMOS technology, which permits the integration of both sensor and processing onto the same substrate. Many research institutes have proposed similar, single-chip architectures [10,11]. However, as

technology advances, transistor sizes become smaller, and hence more processing and logic can be integrated onto the system-on-chip. It is also known that, according to Moore's law, logic speed will also increase with newer, smaller technologies [12]. Yet, the drawback of such progress is the increase of the memory-processor performance gap [13]. The growing disparity between memory and processor speed and size has always been an impediment to embedded systems design. This is why, in the case of the VISoc chip, data memories could not be embedded onto the IC.

Memory manufacturers have always realised the importance of embedded memory, but it wasn't until very recently that system-on-chip RAM began to appear [14]. The new 1T-SRAM technology can have a significant impact on the design of vision systems-on-chip, which rely heavily on memory for image and data storage. The high density, low power consumption, and high speed characteristics offered by the 1T-SRAM technology will enable the development of the next-generation vision systems-on-chip, such as SmartPupilla [15]. These will not only integrate sensor and processing on the same substrate, but will also have room for embedding megabits of memory; minimising bottlenecks, and reducing inter-chip delay and therefore increasing the global system performance.

SmartPupilla is proposed to integrate a VGA sensor [16], a high performance, programmable image pre-processor [17], a neural processor [18], a RISC core for handling interfaces and decision making, as well as high density embedded memory. It is aimed at serving a wider range of applications than VISoc owing to its enlarged sensing area, added processing power and increased throughput. The vision system-on-chip will be ideal for real-time application fields such as security, automotive, and industrial control.

#### 4.1 The Pupilla Sensor

Pupilla is a CMOS VGA sensor with a 120dB logarithmic response. In current stand-alone implementations (0.35 $\mu$ m technology) the pixel array measures 5.12x3.84 mm<sup>2</sup>, and integrates a 10-bit ADC on-chip. Like Vista, Pupilla permits random addressing of pixels, and selection of the area of interest.

#### 4.2 Parallel Processor

SmartPupilla will comprise a novel parallel architecture specifically dedicated to performing image pre-processing tasks. The architecture is based on a parallel array of 16 identical processing elements, capable of executing DSP operations according to a programmable algorithm.

High performance is achieved by an intelligent DMA

channel which addresses the source image according to one of 25 available addressing modes. The set of modes was chosen to cover the most commonly used image processing algorithms, such as windowing, slicing and correlation. Memory bottlenecks are reduced by detecting overlaps in the source regions to be distributed to the processing array, therefore minimising the need for redundant pixel reads whenever possible. The image pre-processing architecture was prototyped on a Xilinx XCV2000E FPGA, connected to the Vista sensor via external I/O. The prototype achieved data throughputs of up to 667 frames/second at a clock frequency of 50 MHz, using 256x256 pixel frames. Its peak performance was estimated to reach 3.23 GOPS. An equivalent VLSI implementation of the architecture in 0.35 $\mu$ m technology is reported to occupy 12.5mm<sup>2</sup>, excluding memory.

#### 4.3 SoftTOTEM

Like VISoc, SmartPupilla will include an N-node parameterisable artificial neural network, where N is the number of nodes defined at synthesis time. SoftTOTEM is the VHDL equivalent of the TOTEM architecture, and as such, it can be ported to any VLSI technology with ease. A 32-node SoftTOTEM has been prototyped on a Xilinx FPGA, achieving a peak performance of 3.30 GOPS at a clock frequency of 50 MHz. An equivalent VLSI implementation in 0.35 $\mu$ m technology is estimated to occupy 10.1mm<sup>2</sup>, excluding weight memories.

#### 4.4 On-Chip Memories

Memory access often poses data and performance bottlenecks on the embedded system. Real-time processing necessitates rapid data retrieval and write-back in order to achieve the required throughput. The use of external memories for data storage can, and quite often does, constrain the processing speed by extending access delays to allow for inter-chip communication.

An alternative approach to memory distribution which reduces access delays is the integration of data memories into the same chip. This approach, however, is not practically feasible in current, mature technologies; certainly not in 0.35 $\mu$ m processes where a single SRAM bit is estimated to occupy about 60 $\mu$ m<sup>2</sup>. Nevertheless, the integration of large amounts of on-chip memories is achievable if modern state-of-the-art embedded memory architectures are used. Even though SoC-RAM architectures are still in their infancy, their use and availability have a promising future.

#### 4.5 RISC Processor

As in VISoc, SmartPupilla will comprise a 32-bit RISC processor which, together with SoftTOTEM and the Image Pre-Processor, will complete the vision tasks. The processor will also be responsible for handling communication with the external world.

Sensor (Pupilla)	640x480 APS
Frame Rate	130 FPS
Resolution	10-bits
On-Chip Processing	32-bit RISC+DSP 16-node Parallel Pre-Processor
On-Chip Mem.	~ 32 Mbits
Neural Network	N-node (parameterisable)
Clock Freq.	≥ 100 MHz
Performance	100 RISC MIPS 6.46 GOPS Parallel Processor 6.60 GOPS Neural Processor
Power	≤ 2 W
Process	0.18 μm or 0.13 μm
Size	~ 150 mm <sup>2</sup>

Table 2: SmartPupilla Characteristics

## 5 CHALLENGES AND FEASIBILITY

Despite their small size, miniature vision systems are not trivial to design and test. Their development, in all its phases from conceptual to physical implementation cannot go unhindered. This is often due to technological and financial constraints. Several trade-offs must be dealt with, and design decisions must sometimes be taken at the conceptual phase of the project as it evolves. Apart from architectural decisions, vision SoC designers are often faced with the open problems of mixed-signal noise, memory distribution, and technological constraints.

### 5.1 Mixed-Signal Noise

Mixing analogue and digital circuits on the same substrate gives rise to the problem of what is known as mixed-signal noise, or crosstalk [19]. This is caused by the injection of noise by the digital components into sensitive analogue circuits, and in the case of vision systems, the analogue circuits make up the image sensor, and hence any mixed-signal noise can cause serious degradation to image sensing quality. In VISoc, this problem had been tackled by introducing a ‘*guardring*’. The guardring is a reverse-biased PN junction which creates a depletion region around the analogue components. The depletion region then acts as a shield from digital noise.

### 5.2 Memory Distribution

In the past, memory distribution had been constrained by fabrication processes. Due to their large size, data memories have often been kept external to systems-on-chip, whilst internal caches helped improve the performance and reduce inter-chip delays. However, as modern-day real-time applications demand high performance and data throughput, vision systems require a significantly large amount of memory to store image data. In the particular case of parallel processors, memory

distribution becomes a challenge. Keeping memories external increases the cost of packaging as the number of I/O pins increases, while speed decreases as a result of inter-chip delay. However, future vision systems-on-chip can benefit from the advances in embedded RAM technology. Embedding memories onto the same chip will increase system speed and performance, and decrease the cost of packaging as the number of I/O pins on the chip is reduced.

### 5.3 Sensor/Logic Technology Conflict

For improved spectral response [20], CMOS imaging sensors must be implemented at reasonably mature processes, such as 0.35μm or 0.25μm, powered by 3.3V or 2.5V supplies, respectively. Newer, smaller technologies, however, can degrade the spectral response of CMOS sensors. Lower voltage supplies are often insufficient to provide a base for detectable analogue swings corresponding to incident light, and this results in the reduction of the sensor’s dynamic range. This constraint conflicts with the technological advances of processing logic, whose performance improves vastly with newer, smaller fabrication processes. To overcome this conflict, novel technologies offer two possible solutions. The first solution is to use Multi-Chip Module technology [21], which can be used to separate the sensor from the rest of the logic, yet maintain the entire system on a single package. MCM technology opens up the opportunity to use different fabrication processes for logic and sensing, and also reduces mixed-signal noise as it separates the analogue and digital substrates. The alternative solution relies on the use of fabrication processes which permit the use of different voltage supplies on the same chip, despite it being built on a small voltage supply foundation. An example of this is TSMC’s 0.18μm process, which is based on a 1.8V supply, and offers 3.3V and 2.5V supplies for mixed-signal cores [22].

Once these design challenges are resolved, the development of miniaturised vision systems-on-chip can proceed. Assuming the use of a 0.18μm fabrication process, with multiple voltage supplies for mixed-signal cores, and embedded 1T-SRAM, SmartPupilla is estimated to occupy a mere 152mm<sup>2</sup>.

Block	Area
Sensor	30 mm <sup>2</sup>
Parallel Processor	3 mm <sup>2</sup>
Neural Processor	3 mm <sup>2</sup>
RISC+DSP	1 mm <sup>2</sup>
Embedded Memory	115 mm <sup>2</sup>
<b>TOTAL</b>	<b>152 mm<sup>2</sup></b>

Table 3: SmartPupilla’s Area Parameters

## 6 CONCLUSION

The paper has discussed the feasibility of miniaturised vision systems, with particular reference to existing and proposed vision systems-on-chip. Small size, light-weight and limited power consumption make these devices unobtrusive; a factor of paramount importance in embedded systems in which the computing functions must be hidden from the final user and communication bandwidth must be minimised. Cost is also continuously reduced as component manufacturers take advantage of price reductions in silicon circuits, as well as technological progresses in embedded memories and mixed-signal design. These implementations of “pervasive computing” concepts are bound to have a huge application potential and wide-ranging implications in control, security, and many aspects of the automotive field. The programmability of such architectures ensures their suitability for general-purpose vision applications, as

opposed to custom, application-specific ICs (ASICs), that are often hard-wired to perform a limited set of tasks related to the particular target application.

The next-generation single-chip smart cameras of the type described here are likely to replace existing devices based on separate sensors and processors, and extend first-generation vision systems-on-chip to include more processing power and embedded memory. This evolution, however, is going to be progressive and will require a number of years, particularly for the sake of cost-effectiveness as the novel technologies described in this paper are still in their infancy and will require a period of time to become available at reasonable, affordable prices. There is no technological barrier to the implementation of miniaturised vision systems. The only remaining challenge to be tackled is that of development costs.

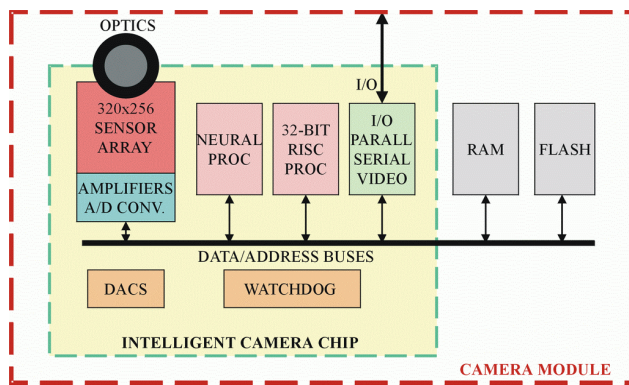
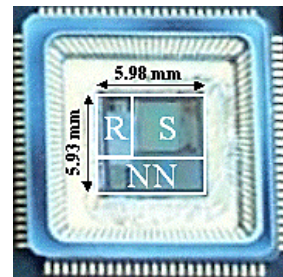


Figure 1: The VISoc Architecture



R = RISC Processor  
S = CMOS Sensor  
NN = Neural Network

Figure 2: The VISoc Die

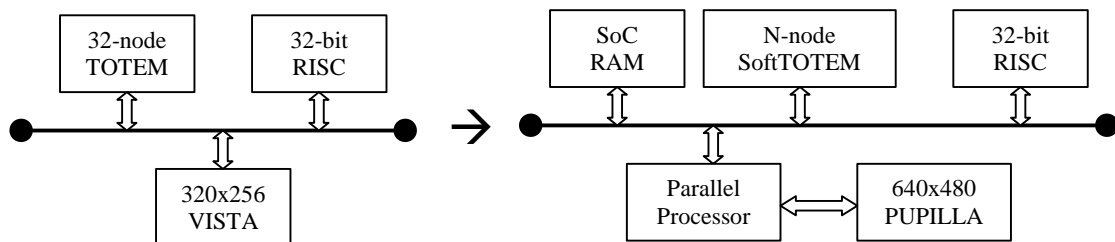


Figure 3: From VISoc to SmartPupilla

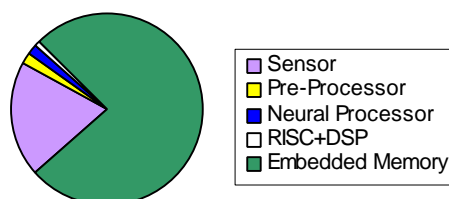


Figure 4: Area Distribution in SmartPupilla

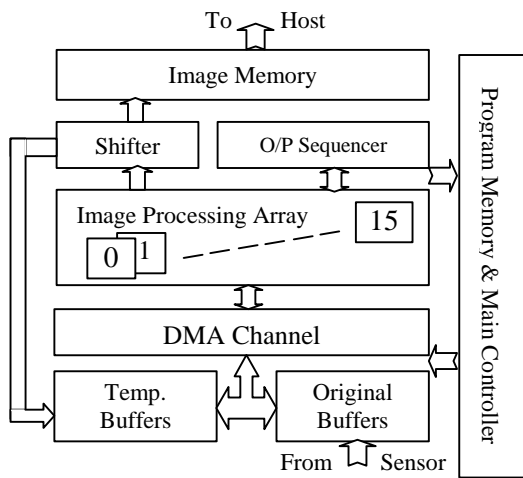


Figure 5: Image Pre-Processor

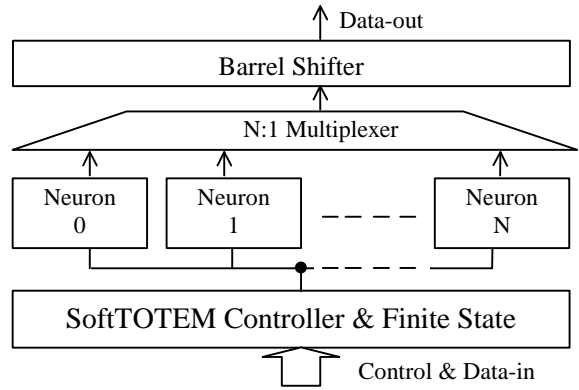


Figure 6: SoftTOTEM

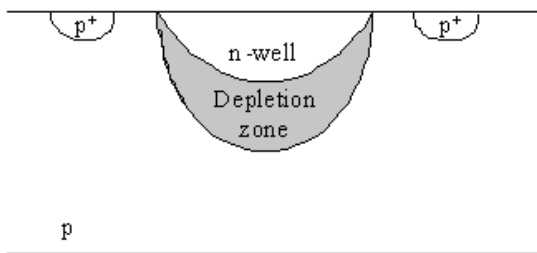


Figure 7: Guardring

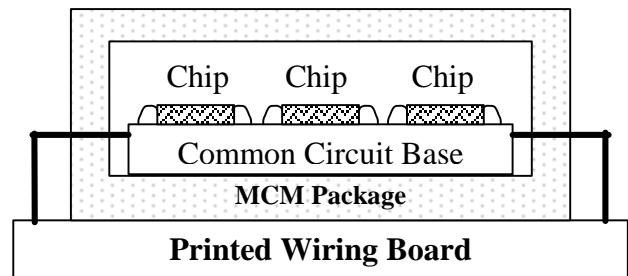


Figure 8: Multi-Chip Module Technology

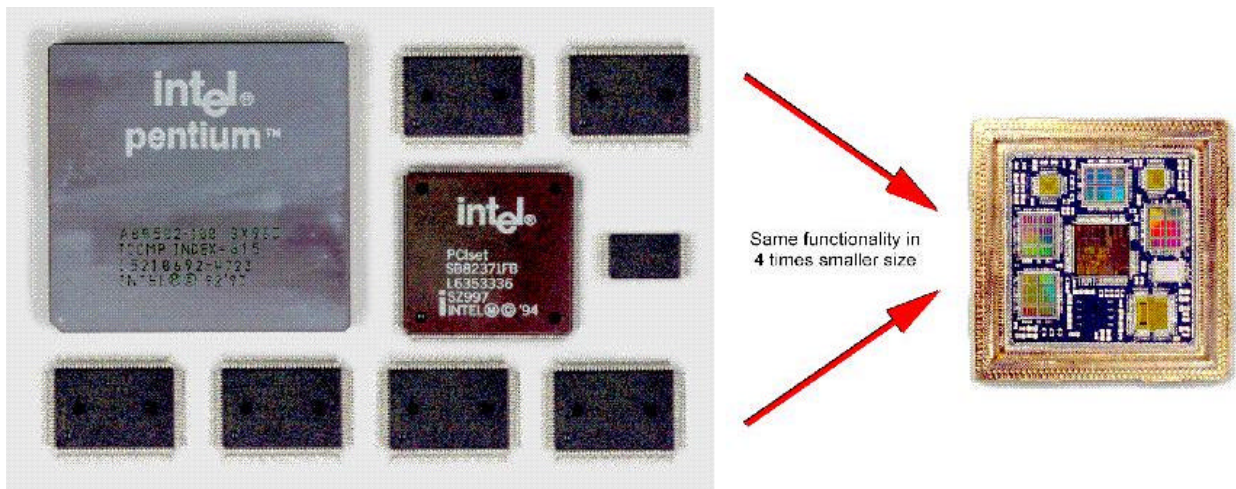


Figure 9: Example MCM Die [23]

## ACKNOWLEDGEMENTS

This work is funded by the European Commission's Marie Curie Host Fellowship contract number HPMT-CT-199-00055 with NeuriCam S.p.A., Italy.

## REFERENCES

- [1] F. Paillet and D. Mercier, "Design Solutions and Techniques for Vision System On a Chip and Fine-grain Parallelism Circuit Integration", in 13th Annual IEEE International ASIC/SOC Conference, September 2000.
- [2] TSMC, Image Sensing Technology, Online Information, [www.tsmc.com](http://www.tsmc.com)
- [3] C. Rowen, "Automated Processor Generation for System-on-Chip", in Proc. of 27<sup>th</sup> European Solid-State Circuits Conference, pp.29-34, Villach, Austria, September 2001.
- [4] R. Habets, "Machine vision for automatic inspection in the food processing industry", in ", Proceedings of ACIVS 2002 (Advanced Concepts for Intelligent Vision Systems), pp.1-5, Ghent, Belgium, September 2002.
- [5] NeuriCam S.p.A, "EtherCam Intelligent Camera", Product Information, 2002.
- [6] Wai-Chi Fang, "Low-Power, Fast Machine Vision System on a Single IC Chip", February 2000, <http://www.nasatech.com/Briefs/Feb00/NPO20449.html>
- [7] U Ramacher *et al*; A 53-GOPS Programmable Vision Processor For Processing, Coding-Decoding And Synthesizing of Images, Proceedings of the 27th European Solid-State Circuits Conference, pp.160-163, Villach, Austria, September 2001.
- [8] A Abbo *et al*; A Low-Power Parallel Processor IC for Digital Video Cameras, Proceedings of the 27th European Solid-State Circuits Conference, pp.164-167, Villach, Austria, September 2001.
- [9] L. Albani *et al*, "VISoc: a Smart Camera SoC", in Proc. of 28<sup>th</sup> European Solid-State Circuits Conference, pp.367-370, Firenze, Italy, September 2002.
- [10] Wai-Chi Fang, "A smart vision system-on-a-chip design based on programmable neural processor integrated with active pixel sensor", Proceedings of The IEEE International Symposium on Circuits and Systems, Vol.2, pp.128-31, Geneva, Switzerland, May 2000.
- [11] A. Abbo and Richard P. Kleihorst, "A Programmable Smart Camera Architecture", Proceedings of ACIVS 2002 (Advanced Concepts for Intelligent Vision Systems), pp.6-13, Ghent, Belgium, September 2002.
- [12] Stefan Rusu, "Trends and Challenges in VLSI Technology Scaling Towards 100nm", in Proc. of 27<sup>th</sup> European Solid-State Circuits Conference, pp.23-25, Villach, Austria, September 2001.
- [13] N. Mahapatra and B. Venkatrao, "The Processor-Memory bottleneck: Problems and Solutions", ACM Crossroads, Computer Architecture, Spring 1999 - 5.3.
- [14] W. Leung, F. Hsu and M. Jones, "The Ideal SoC Memory: 1T-SRAM", Monolithic System Technology Inc., Technical Article, [www.mosys.com](http://www.mosys.com)
- [15] S McBader *et al*; "SmartPupilla: A Complete Vision System-on-a-Chip", in ", Proceedings of ACIVS 2002 (Advanced Concepts for Intelligent Vision Systems), pp.224-228, Ghent, Belgium, September 2002.
- [16] NeuriCam S.p.A, "NC1802 Pupilla CMOS VGA Sensor", Datasheet, 2002.
- [17] S McBader & P Lee, "A Programmable Image Signal Processing Architecture for Embedded Vision Systems", Proc. 14th Int. Conf. on Digital Signal Processing, Santorini, pp.1269-1272, Greece, July 2002.
- [18] S McBader, *et al*; "SoftTOTEM: An FPGA Implementation of the TOTEM Parallel Processor", Proc. 12<sup>th</sup> Int. Conf. On Field Programmable Logic and Applications (Lecture Notes in Computer Science LNCS 2438), pp.1084-1087, Montpellier, France, September 2002.
- [19] W. Sansen, "SoC Design from a Mixed Signal Perspective", in Proc. of 27<sup>th</sup> European Solid-State Circuits Conference, pp.2-5, Villach, Austria, September 2001.
- [20] Markus Loose, "A Self Calibrating CMOS Image Sensor with Logarithmic Response", PhD Thesis, Institut Fur Hochenergiephysik, Universitat Heidelberg, October 1999.
- [21] Hitachi Semiconductors, MCM technology information.
- [22] TSMC, 0.18-Micron Technology, Online Information, [www.tsmc.com](http://www.tsmc.com)
- [23] EuroPractice, "A Pentium Processor MCM for Embedded Systems", <http://www.europractice-hdp.com/products/pentium.html>