Efficient Home Video Surveillance Platform

Th. Zahariadis[1], K. Gröneberg[2] and TK Chiew[3]

[1.] Algosystems S.A.
206 Syggrou Av., Athens, GR 17673, Greece
Phone: (+30) 210 9548 040 Fax: (+30) 210 9548 099 E-mail: zahariad@ellemedia.com

[2.] Fraunhofer Institut fur Nachrichtentechnik (FhG/HHI)
37 Einsteinufer Str., Berlin, 10587, Germany

[3.] Institute For Infocomm Research
21 Heng Mui Keng Terrace, 119613, Singapore

Keywords: SVC, surveillance, home networking

Abstract - Resent market research shows that an increasing amount of consumers are interested in multimedia Smart Home applications. Moreover, home security/surveillance is considered as among the enabling applications. In this paper, we describe an approach to capture these markets, by implementing an innovative Audio/Video platform, able to pioneer in the near- to mid- market segments. Moreover, we propose a scalable surveillance platform.

1. INTRODUCTION

Resent market research shows that an increasing amount of consumers are interested in multimedia Smart Home applications, like sharing music, pictures, and videos among digital electronics, PCs, and mobile devices. Subsequently, this will increase the home networking equipment demand, leading world market revenues up to 5.5 billion dollars by 2006 [1]. The Global Market Forecast reveals that multimedia subscribers will grow to 26 million in 2008. Many forecasts also illustrate that Video on Demand (VoD) and home security/surveillance have moved from add-on services to early deployment services. In this paper, we describe an approach of ASTRALS[1] project to capture these markets, by implementing an innovative Residential Gateway (RG) Audio/Video (A/V) extension platform, able to pioneer in the near- to mid- market segments. Moreover, we propose a scalable A/V surveillance platform, which is currently under implementation.

ASTRALS (Audio-visual STReaming plAtform for domestic Leisure and Security) is focused on scalable, A/V encoding, transcoding, storage and distribution in existing households via streaming-optimised wireless links. The motivation is to implement scalable solutions, which will enable a new beam of innovative A/V products and services, including personalised, network-aware video adaptation and distribution in multiple heterogeneous terminals (from low-cost PDAs to high-end home cinemas) and intelligent surveillance, utilising a state of the art in-home network.

The paper is structured as follows. Initially, it presents the envisaged home network architecture and the surveillance application requirements. Then it describes the video coding requirements and outlines the necessary main A/V sub-system capabilities. However, the design decision of the system architecture is not based only on technical facts. Techno-economical criteria have also been taken into account to provide an abstract system architecture and specification.

2. HOME SURVEILLANCE ARCHITECTURE

We envisage a multimedia-centred network architecture able to capture, encode, process and distribute efficiently A/V streams. The whole video processing and wireless transmission process is co-ordinated by the RG (Fig. 1), which interconnects access and indoor networks and provides enhancements for A/V encoding, transcoding and storage of user/ services profiles and content. Without excluding Data (Ethernet, Wireless LAN, UWB), Analogue (Home Cable & RF) or Digital (Firewire, USB) network interfaces, we assume a streaming-optimised, wireless network as the major distribution medium. Through multimodal devices and/or 4G terminals, a new spectrum of in-home networked applications, ranging from simple Internet access over "don't care" communication links to advanced creation, transcoding and access on streamed/ stored A/V media from anywhere in/out the home environment will be provided.

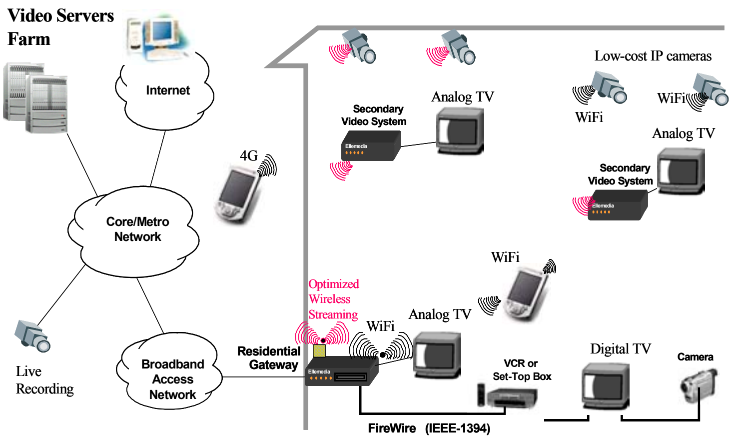

Fig. 1: Envisaged Network Architecture

Various market studies have shown that although consumer views about Smart Homes are fairly mixed, there is a significant level of underlying interest in the concept of home security. In a sample of 1,000 households, security features emerged as the most popular aspect of living in a Smart Home, 70% of respondents

[1] ASTRALS is an Information Society Technologies (IST) project co-funded under EU's contract FP6-IST-0028097

agreed that they would really value the safety and security features a Smart Home, having the benefits of remote access also wide appeal (59%) [1]. Home surveillance mainly aims at enhanced security for home users. Other important applications are healthcare, surveillance of elderly people and baby-sitting.

Most surveillance systems however, face a number of issues, concerning the adequate system specification and the adoption of a specific technology. Issues that should be considered include: a) Image quality, b) Resolution, c)Frame Rate, d) Coverage, e) Environmental Conditions, f) System Control, g) Storage, h) Cost.

2.1 Video Content Analysis

Most video surveillance systems provide Digital Video Recorder (DVR) functionality, which is used for video recording and watched later by a human. We propose to utilise Video Content Analysis (VCA), which extracts high-level information from the video input and enables online event alert and fast retrieval of event related video. According to the processing and coding order, there are two options for the interface between the Home Surveillance and the A/V platform.

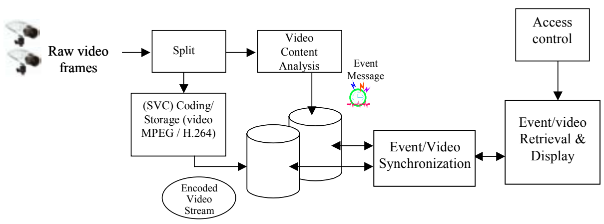

Fig. 2: Independent encoding and VCA processes

In the first case (Fig. 2), raw video frames are duplicated and forwarded independently to the Video Encoding and VCA modules. The two modules operate independently and store their results (encoded video streams and event messages respectively) as separate entries at the storage devices (local disk). The major advantage of this method is flexibility. By having the video encoding and VCA modules independent, they can be atomic modules replaced and/or upgraded at any time. The disadvantage of the method is that event detection messages and stored video streams need to be synchronized before they are sent to the display (draw the event on the video stream/frames).

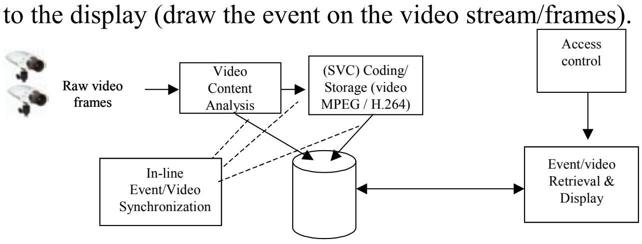

Fig. 3: Encoding follows the VCA process

In the second case (Fig. 3), the VCA process is firstly applied to the raw video, which is sequentially encoded by the encoder module. Synchronization in this case takes place in-line on a per frames basis. The major advantage of this method is that video streams and events can be synchronized on the frames output from VCA module. On the other hand, depending on the VCA computational complexity, the VCA output is usually low-frame-rate video (frames), e.g. 8-16 fps. So the stored video after encoding is not in full frame rate, it could be also variable frame rate.

3. Scalable Video Encoding

In order to maximize video portability, scalability and error resilience across a number of terminals, we have selected Scalable Video Coding (SVC) as the major encoding standard. The SVC intents to create a standard for efficient video compression that provides bit streams scalable in frame rate, resolution and SNR quality. The SVC extension is built on H.264 / MPEG-4 AVC and re-uses most of its innovative components [3]. Initially, SVC generates a backwards-compatible H.264/MPEG-4 AVC compliant base layer and one or several enhancement layer(s). The base layer bit stream corresponds to a minimum quality, frame rate, and resolution (e.g., QCIF video), and the enhancement layer bit streams represent the additional information needed to improve the same video with gradually increasing quality and/or resolution (e.g., CIF) and/or frame rate.

3.1 Processing Power

The SVC encoder processing requirements depend on several parameters and scale with image size and frame rate of the input video. On the other hand, the compression ratio and the image quality are limited by the available processing power. Last but not least, the requirements strongly depend on the desired number of scalability dimensions and levels. There is no upper limit to the processing power, because any increase could be used to implement new features, such as additional scalability levels or dimensions, or improve image quality or reduce data rate or both.

At this moment, only the reference software for the scalable extensions of H.264/MPEG-4,Part-10 is available, which is not optimized regarding processing speed. First tests with different encoder settings have been carried out on different sequences (Fig. 4). The encoder was run on a Xeon CPU clocked at 1.7 GHz. As Motion Estimation (ME) is the most demanding module within the encoder, we tried and restricted motion estimation to the base layer (red boxed markers in Fig. 5).

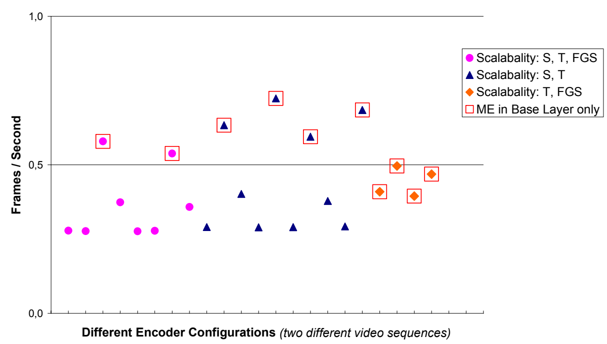

Fig. 4: SVC Encoder Benchmarks (non-optimized Software)

Still the encoder is much too slow. Code optimization alone will hardly speed up the encoder by a factor of about 30, which would be required for real-time processing of CIF format video. As a reference, an optimized single layer H.264/MPEG4-AVC encoder, running on an Intel Pentium 4 Xeon 64-bit CPU clocked at 3.6 GHz, is able to encode CIF images at a frame rate of 25 frames/s. Encoding with e.g. two spatial scalability levels and two quality scalability levels will need much more processing power and DSP optimization may be required

3.2 A/V Storage

Usually, surveillance videos are required to be stored for at least a week before they are overwritten. However, storing even compressed video streams is very much space consuming. Just for reference, one day of MPEG 1 (CIF: 352x288, 25fps) video stream requests approximately a storage area of 24GBytes on HDD, whereas the same stream, SVC encoded at an average data rate of 350 kbit/s, uses less than 4GBytes of HDD space per day. If a camera faces no motion at all for most of the time, it will be much less. In ASTRALS, the SVC erosion feature will be utilised. SVC erosion feature is based on the time stamp of the video files. Once the HDD is nearly full, the 10% oldest files will be deleted.

4. A/V Platform Requirements Analysis

The proposed A/V platform aims to capture a large market share of residential A/V surveillance systems. The main requirements for the video-processing subsystem, as they are also reflected in the previous sections may be summarized in the following:

a) *Adequate processing-power.* A/V real-time AVC/ SVC encoding and transcoding/ transrating have extremely demanding nature in terms of processing-power.
b) *Compact/limited space.* It should be quite compact and should dissipate limited power, so that it will be either integrated with the RG into a single enclosure or be available as a small add-on system.
c) *Storage.* The A/V subsystem should be able to store encoded streams locally on the sub-system or centrally on the main RG system.
d) *Cost-effectiveness.* In respect to commercial viability, it is very important that the final selection leads into a cost effective solution.
e) *Openness & flexibility.* The term "openness" means that the provided solution should allow future upgrades.

Having all these basics in mind, we proceed with a further analysis of the A/V platform requirements analysis.

4.1 Embedded Processor

The embedded processor is a key component in the design of the A/V platform. ASIC and ASSPs are not considered as they have limited flexibility. Digital Signal Processors (DSPs) are ideal for specialised processing demanding applications like AVC/ SVC encoding and transcoding/transrating applications. Moreover, DSPs increasingly have instructions or peripherals that offer special advantages in certain applications. For example, motion control is receiving a great deal of attention from many multimedia DSP vendors. On the other hand General purpose processors (GPP) offer great programmability and therefore flexibility. They may not be as efficient in signal processing functions as DSPs, however modern GPPs are very powerful and quite cost effective and include dedicated processing engines and A/V extensions. Finally, there is huge amount of open source code that can be reused in an architecture that features a GPP.

Taken into account the requirements mentioned in the previous sections, we have decided that the most viable solution is either a GPP or the hybrid GPP/DSP solution. We considered the following high-end GPP alternatives:

a) *Pentium-M* and *Celeron-M* devices by Intel
b) *mobile-Athlon* and *Turion* by AMD
c) *Luke*, *Eden, C3* and *C7* by VIA.

New processors have been announced (i.e. dual-core). However, due to strict workplan, we have to limit our considerations on the above systems. It is quite difficult to have a straight-forward comparison of the aforementioned processors, since they differ in many aspects. Furthermore, the overall system performance is highly influenced by the accompanying chipset. It is clear from all benchmarking tests that the Intel Pentium-M and AMD Turion-64 hold the top, in terms of performance. In order to compare the Pentium-M and Turion-64 devices, we've discovered through several, often contrasting, benchmarking reports, that both devices are more-or-less equal in terms of performance, while Pentium-M outperforms in terms of power consumption and compactness. Furthermore, the dominance of the Pentium-M devices in the notebook market, also suggests that Pentium-M pave the way.

The VIA processors fall well behind, in terms of processing power, however they outperform in terms of power-consumption and compactness. Additionally, they incorporate a hardware MPEG2/4 accelerator, which however will not be utilised. Another factor is the power consumption. As it is shown in the following graph, VIA processors consume considerable less power than similar Intel chips. However, as the A/V subsystem will not be applied to a mobile device, we consider power consumption as having lower importance.

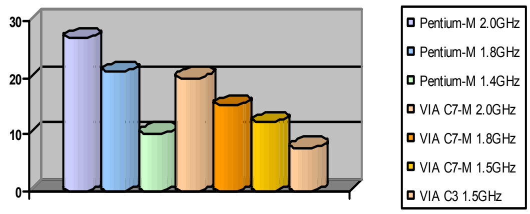

Fig. 5: Total power for Intel & VIA processors [3]

Processor	Max Clock Speed	Front-Side-Bus speed	Total Power	Size (mm^2)	Price (small quantities)
Intel Pentium-M	2.26GHz	533 MHz	27W	35x35	250-350 €
VIA C3/ Eden	1.5GHz	400 MHz	7.5W	35x35	80-140 €
VIA C7-M	2GHz	800 MHz	20W	21x21	100-150 €

Table 1. Processors Comparison

Table 1 presents the basic characteristics of the main candidate processors. We concluded that a Pentium-M based platform is currently the most powerful architecture. There is no evidence that the processing power provided by a VIA C3/Eden or Luke processor will be adequate, while, samples of VIA C7 chips are not widely available as yet. The use of a DSP co-processor (e.g. TI DaVinci) is also a good option, which could enable lower-performing platforms to support the particular video applications. However, this would result in DSP specific programming, which would limit the benefit from SVC encoding.

We concluded that a Pentium-M based platform is the most powerful architecture, for the time being. There is no evidence that the processing power provided by a VIA C3/Eden or Luke processor will be adequate, while, samples of VIA C7 chips are not widely available as yet. The use of a DSP co-processor or a DSP based platform (e.g. the DaVinci) is also a very good option, which could enable lower-performing platforms to support the particular video applications. However, this would result in DSP specific programming, which would limit the benefit from SVC encoding.

Thus, taking into account the processing-power, ease of development and cost effectiveness, the most appealing solution, in the terms of the project, is to utilize a compact, but as powerful as possible, Pentium-M architecture that will enable us to develop a space-restricted A/V subsystem. This subsystem may require the use of a DSP co-processor in order to support the various video applications. However, due to the fact that, at this initial point of the project, we do not have a clear view of the exact processing requirements, we'll design the A/V subsystem with the option to address a DSP co-processor at a later stage via the system PCI interface.

5. A/V Platform Specifications

Taking into account the considerations of the previous section we conclude on the following system architecture and hardware specifications.

The system architecture is shown in Fig. 6. The system is based on the Intel Processor-M and the accompanying chipsets Intel 828GME (82852GME) and ICH4 [4]. The chipset is interconnected utilising FSB 400/533MHz and Hub Interfaces respectively. The 828GME features an SDRAM socket, where one DDR333 memory module will be connected. Based on the processor up to 1GB SDRAM will be connected. Optionally a Microphone and/or Audio Out interface and a VGA and/or a TV output interface would be provided.

The system also features the Intel ICH4 where a number of interfaces will be connected. First of all, the FWH provides a flash Bios with 8Mbits. Then a 33MHz PCI Bus offers at least 1-2 PCI Bus slots and 1 miniPCI interface. One of the PCI slots will be allocated for the DSP option, while the other PCI and the miniPCI will be utilised for a number of subsystems e.g. a miniPCI capture card or a WiFi or a IEEE 1394 daughter board. Moreover, at least 2 x USB 2.0 interfaces are provided to support USB cameras. The system interfaces the RG and/or the Wireless card via a 10/100 Base-T Ethernet interface. The HDD and/or the Compact Flash are connected to an ATA/100 IDE slot. Optionally, an RS232 and/or an IrDA interface may be attached for control purposes.

In Fig. 6 a possible configuration is also shown. The configuration features a miniPCI Video Capture card with two BNC connectors where two monitor cameras may be attached. Moreover, a DSP daughterboard for extra signal processing power is connected to one PCI interface. Efficient transcoding and scalable encoding will be provided via efficient integration between the Video Grabber or the MPEG2/MPEG4 Video source, the Wireless Manager and the Profile manager (Fig. 7).

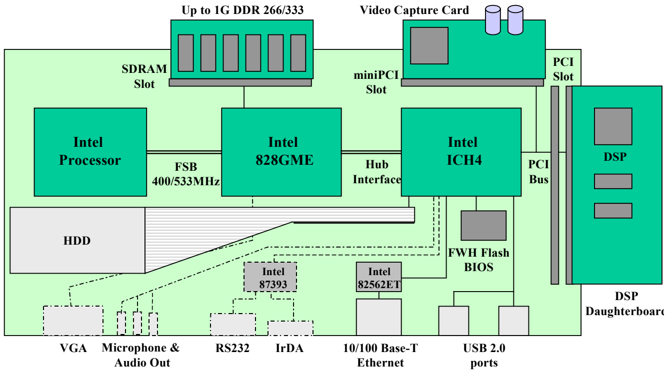

Fig. 6: A/V subsystem configuration

Video streams may feed directly the SVC Encoder or Transcoding/transrating utilities or be temporarily stored in a Video Storage database. The Video Content Analysis will operate in parallel. The encoders/transcoders outcome video streams will be either stored at the Encoded Video Storage database or forwarded to the Video Server and transmitted to the Wireless subsystem. In case of SVC encoder additional SVC signalling will be provided.

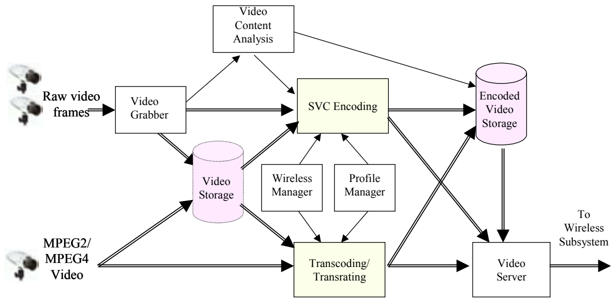

Fig. 7: A/V Software Interfaces

4. CONCLUSION

Recent market research shows that home security/ surveillance is considered as among the enabling smart home applications. In this paper, we described an approach to capture these markets, by implementing an innovative Audio/Video platform, able to pioneer in the near- to mid-market segments.

Acknowledgment

This publication is based on work performed in the framework of the Project ASTRALS, which is partially funded by the European Community. The authors would like to acknowledge the contributions of colleagues from: Algosystems, STMicrolectronics, Thomson, Mitsubishi Electric, Telefonica I+D, Provision Communications, Fraunhofer HHI, Institute for Infocomm Research.

References

[1] Multimedia Research Group, Inc, "IP/Broadband Video Market Tracking Service," 2004

[2] H. Schwart, D. Marpe, T. Wiegand, "Basic concepts for supporting spatial and SNR scalability in the scalable H.264/MPEG4 AVC Extension," 12th IWSSIP, Chalkis, Greece, 22-24 September 2005, pp.10-14

[3] http://www.via.com.tw/

[4] http://www.intel.com

[5] http://www.ist-astrals.org