

Guest editorial: special issue on algorithms and architectures for real-time image and video enhancement

Sergio Saponara · Giovanni Ramponi ·
Stefano Marsi · Gerard de Haan · Erwin Bellers

Received: 25 October 2011 / Accepted: 9 November 2011 / Published online: 20 November 2011
© Springer-Verlag 2011

Recent advances in real-time image and video enhancement are enabling innovations in a broad range of applications including biomedicine, intelligent transportation, driving assistance, consumer electronics, telecommunication, robotics and surveillance. These innovations encompass complexity-aware algorithms and new hardware–software (HW–SW) architectures and aims at (1) improved visualization performance; (2) acceleration of processing (real-time instead of off-line computation); and (3) complexity reduction to meet demands involving device size, power consumption, cost of target applications such as battery-powered mobile/wearable devices, or low-cost large volume markets.

This special issue presents nine papers covering different real-time algorithms and cost-efficient architectures for several image/video enhancement techniques. The accepted papers are from different international institutions located in North and South America, Europe and Asia.

The research on image/video enhancement used to be mainly focused on multimedia, consumer or telecom

applications. However, this special issue demonstrates the growing interest for image/video enhancement techniques to fields such as biomedicine (capsule endoscopy and neuroscience test), robotics for automation in agriculture, automotive driving assistance, and aerial surveillance. The discussed techniques include object tracking, image and video compression, edge extraction/detection for image analysis, anomaly detection, lighting conditions improvement, and contrast enhancement.

The first paper by Khan et al. presents a subsampling-based image compressor for capsule endoscopic system which is aimed at reducing the chip area and power consumption, while maintaining an acceptable video quality. A low-complexity algorithm, suitable for VLSI implementation, is developed around some special features of endoscopic images and consists of a differential pulse code modulation followed by Golomb–Rice coding. An image corner clipping algorithm is also presented. The reconstructed images are verified by medical doctors for acceptability. Compared to other transform-based algorithms targeted to capsule endoscopy, the proposed raster-scan-based scheme performs very strongly with a compression ratio of 80% and a very high reconstruction PSNR (over 45 dB).

The second paper by Armato et al. also deals with biomedical-related applications. This work is focused on exploring and comparing several photometric normalization techniques to improve eye gaze tracking (EGT) systems during light changes. EGT is developed for scientific exploration in controlled environments where it is used in ophthalmology, neurology, psychology, and related areas to study oculomotor characteristics and abnormalities, and their relation to cognition and mental states. The illumination is one of the most restrictive limitations in EGT, due to the problem of pupil center estimation during illumination

S. Saponara (✉)
Dip. Ingegneria della Informazione, Università di Pisa,
via G. Caruso 16, 56122 Pisa, Italy
e-mail: sergio.saponara@iet.unipi.it

G. Ramponi · S. Marsi
Dip. di Ingegneria Industriale e dell'Informazione,
Università di Trieste, via A. Valerio 10, 34127 Trieste, Italy

G. de Haan
Philips Research Laboratories, High Tech Campus 36,
5656AE Eindhoven, The Netherlands

E. Bellers
Trident Microsystems, Inc., 1170 Kifer Rd, Sunnyvale,
CA 94086-5303, USA

changes. A new wearable and wireless tracking system, called HATCAM, is used for testing different techniques in terms of real-time capability, eye tracking and pupil area detection. Embedding real-time image enhancement into the HATCAM can make it an innovative and robust system for eye tracking in different lighting conditions, i.e. darkness, sunlight, indoor and outdoor environments.

The third paper by Xuming Chen et al. presents an interesting application of edge detection/correction plus image recognition for automation in agriculture: ripe tomato picking by a robot using a video system for recognition and localization. The key contribution of this work, based on a three-step low-complexity algorithm, is the real-time recognition and localization of ripe tomatoes in an uncontrolled environment, where they can be covered by foliage, stems and unripe tomatoes.

The fourth paper by Acito et al. deals with a complexity-aware algorithm architecture for real-time enhancement of local anomalies in hyper-spectral images. Anomaly detection (AD) from remotely sensed multi-hyper-spectral images is a powerful tool in many applications, such as strategic surveillance and search and rescue operations where an airborne hyper-spectral sensor searches a wide area to identify regions that may contain potential targets. While this procedure is mostly automated, an onboard operator is generally assigned to examine in real-time the AD output and select the regions of interest to be sent for cueing. Due to the real-time enhancement of local anomalies in images, the proposed technique facilitates the decision-making process. The results on real hyper-spectral images are also presented.

The following five papers deal with the design of innovative architectures for real-time implementation of image/video enhancement tasks.

The paper by Biswal et al. presents a pipelined architecture to accelerate affine transforms used in various high-speed applications such as optical quadrature microscopy, image stabilization in digital cameras, and image registration. The architecture is mapped into a field programmable gate array (FPGA), and the results show that the proposed algorithm is almost four times faster than the conventional algorithm, with a performance of up to 540 fps with images of $1,920 \times 1,080$ pixels, while retaining the image quality.

As discussed in the papers by Marsi et al. and by Happe et al., FPGAs are evolving as complex HW–SW platforms provide a configurable HW logic as well as embedded microprocessor cores for a SW-programmability solution. In the first work, this platform is successfully used for a real-time video enhancement algorithm based on a modified version of the Retinex approach. A new illumination estimation technique is presented, which allows the user to control the dynamic range of poorly illuminated images and to preserve the visual details. Digital cameras, and new

generation of phones, commercial TV sets and nearly all modern devices for image acquisition and visualization can benefit from this solution. The video enhancement parameters are controlled through the embedded microprocessor, which enables the system to modify its behavior according to the characteristics of the input images, and to use information concerning the surrounding light conditions.

The work by Happe et al. presents a video tracking application modeled on top of a framework for implementing Sequential Monte Carlo methods on CPU/FPGA-based systems. Based on a multi-threaded programming model, the proposed framework allows for an easy design space exploration with respect to the HW/SW partitioning. Additionally, the application can adaptively switch between several partitioning states during run-time, in order to react to changing input data and performance requirements. To evaluate its performance and area requirements, the authors demonstrate the application and the framework on a real-life video tracking case study and show that partial reconfiguration can be effectively and transparently used for realizing adaptive real-time HW/SW systems.

When real-time and power consumption are both key issues, as in wearable or mobile battery-powered systems, an effective implementation solution is represented by the realization of a system-on-chip (SoC) using submicron CMOS technology. Not only the computation core but also the on-chip communication infrastructure and the main memory hierarchy has to be optimized, since in video systems complexity and power consumption are often dominated by data storage, transfer rate, the relevant memory size, and access frequency.

To this end, Saponara et al. present a multi-processor SoC architecture for real-time, low-power image and video enhancement applications. Different from other state-of-the-art parallel architectures, the proposed solution is composed of heterogeneous tiles. The tiles have computational and memory capabilities, support different algorithmic classes and are connected by a novel network-on-chip (NoC) infrastructure. The proposed packet-switched data transfer scheme avoids communication bottlenecks when more tiles are working concurrently. The functional performances of the NoC-based multi-processor architecture are assessed by presenting the achieved results when the platform is programmed to support different enhancement algorithms for still images or videos such as contrast enhancement, dynamic range luminance correction, image and video compression, artifacts and noise removal. Consumer devices and automotive driving assistance applications are considered as case studies. Implementation results in 65 nm CMOS technology are proposed. The SoC complexity amounts to 1 million logic gates and 19 Mb of on chip SRAM memory. Running at 400 MHz, the MPSoC ensures real-time processing up to 30 VGA frames per

second with a power consumption of a few hundreds of mW.

The last paper by Zatt et al. presents an optimized motion compensation hardware architecture for the High 4:2:2 profile of H.264/AVC video coding standard. The proposed design focuses on real-time decoding for HDTV $1,920 \times 1,080$ images at 30 fps, with quarter sample accuracy. Multiple sample bit-width and multiple chroma subsampling formats are supported. A novel memory hierarchy is also implemented as a 3-D Cache. It reduces the frame memory access, providing on average reductions by 62% in bandwidth and by 80% in clock cycles. The design is implemented in a Xilinx Virtex-II FPGA, and also in an ASIC with a $0.18 \mu\text{m}$ CMOS technology which occupies 102 K equivalent gates, 56.5 KB of on-chip SRAM in a $3.8 \times 3.4 \text{ mm}^2$ area, and with a power consumption of 130 mW.

In conclusion, the guest editors hope that the selected papers will provide the readers with interesting samples of present research on algorithms and architectures for real-time image and video enhancement in a broad range of applications. They are very grateful to the reviewers who provided valuable comments and suggestions to improve the quality of the accepted papers.

Author Biographies

Sergio Saponara got the Laurea degree, cum laude, and the Ph.D. in electronic engineering from the University of Pisa in 1999 and 2003, respectively. In 2002 he was with IMEC, Leuven (B), as Marie Curie research fellow. Since 2001 he collaborates with Consorzio Pisa Ricerche in Pisa. He is senior researcher at University of Pisa in the field of electronic circuits and systems for telecom, multimedia, automotive applications. He holds the chair of electronic systems for automotive and automation at the Faculty of Engineering. He coauthored more than 150 scientific publications and more than 10 patents or patent applications. Sergio Saponara is also research associate of CNIT and INFN and served as guest editor of special issues on international journals and as program committee member of international conferences.

Giovanni Ramponi is professor of electronics at the University of Trieste. His research interests include nonlinear digital signal processing, enhancement and feature extraction in images and image sequences, image visualization, image quality evaluation. He is the co-inventor of various pending international patents and has published more than 170 papers in international journals, conference proceedings and as book chapters. Prof. Ramponi was from 1997 to 2002 an associate editor of the *IEEE Signal Processing Letters*; from 2002 to 2006 was an associate editor of the *IEEE Transactions on Image Processing* and since 1997 has been an associate editor of the *SPIE*

Journal of Electronic Imaging. He was (co-)chairman of the technical programme of ISISPA-2011, NSIP-2003, IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, and of EUSIPCO-96, Eurasp European Signal Processing Conference. More details on the Web site: <http://www.units.it/ramponi>.

Stefano Marsi received the Dr. Eng degree in electronic engineering (summa cum laude) from the University of Trieste, Italy, in 1990 and the Ph.D. degree from the University of Padova, Italy, in 1994. Since 1995, he has been a researcher in the Department of Electronics, University of Trieste, Trieste, Italy, where he teaches courses in electronic field. His research interests include nonlinear operators for image and video processing and their realization through application specific electronics circuits. He is author or coauthor of more than 50 papers in international journals, proceedings of international conferences or contributions in books. He participated in several international projects and he is the counselor of the local IEEE student branch.

Gerard de Haan received B.Sc., M.Sc. and Ph.D. degrees from Delft University of Technology in 1977, 1979 and 1992, respectively. He joined Philips Research in 1979. Since 1988, he teaches post-academic courses for the Philips Centre for Technical Training at various locations in Europe, Asia and the US. In 2000, he was appointed research fellow in the Video Processing & Analysis group of Philips Research Eindhoven, and full professor at the Eindhoven University of Technology teaching “Video Processing for Multimedia Systems”. He has a particular interest in video/image analysis/processing and computer vision. His work in these areas has resulted in 3 books, 2 book chapters, about 150 scientific papers, more than 100 patents and patent applications, and various commercially available ICs. He received five Best Paper Awards, the Gilles Holst Award, the IEEE Chester Sall Award, bronze, silver and gold patent medals, while his work on motion received the EISA European Video Innovation Award, and the Wall Street Journal Business Innovation Award. Gerard de Haan serves in the program committees of various international conferences.

Erwin Bellers received his M.Sc. degree in Electrical Engineering with distinction from University of Twente in the Netherlands in 1993 and his Ph.D. degree on de-interlacing from Delft University of Technology in 2000. Dr. Bellers joined Philips Research in Eindhoven, the Netherlands, in 1993 as a researcher and mainly conducted research in video enhancement, de-interlacing and frame-rate conversion. In 2000 he joined Philips Research USA in New York as a senior scientist where he led a research-team focusing on spatial resolution enhancement, which was commercialized as ‘PixelPlus’. In 2002 he moved to Silicon Valley to join Philips Semiconductors as a senior video architect. Philips Semiconductors became NXP Semiconductors in 2006. Dr. Bellers further innovated on the algorithm and architecture of frame-rate conversion which became a successful product range of NXP semiconductors. With the merger of NXPs TV business and Trident, he moved along and became responsible for video algorithm innovations within Trident. Dr. Bellers was invited as a coauthor for the *Proceedings of the IEEE*, and he won the second prize in the ICCE Outstanding Paper Awards program in 1997 and 2006, and the first prize in 2009. He has published over 50 papers, 1 book, and holds approx. 50 patent applications.