IEEE AVSS 2013 panel discussion

Are smart cameras ready for prime time?

Moderator

Andrea Cavallaro, *Queen Mary University of London*, UK

Participants

Christian Fabre, *CEA LETI*, France

Charles Attwood, *Thales Research and Technology*, UK

Piotr Szczuko, *Gdańsk University of Technology*, Poland

Toygar Akgün, *Aselsan*, Turkey

A couple of definitions …

- Smart Camera
  - optical device for capturing images
    with embedded hardware
    to extract, process and communicate information

- Cognitive & Perceptive Video Systems
  - system of largely distributed and power-savvy devices (smart cameras) that can communicate among each other

Smart cameras: challenges

- Operations
  - multiple cooperating cameras: centralised *vs.* distributed processing
  - scalability and pervasiveness
- Security
  - access to data by third parties
  - manipulation of data by third parties
- Evaluation
  - datasets & benchmarks
  - performance levels
- Societal issues
  - privacy: can smart cameras help addressing this issue?

The AVSS 2013 panel: topics

- Existing technologies, barriers & trends
  (hardware & software)
- Current (successful) applications
  (security, safety, counting, monitoring, ...)
- Potential (disruptive) applications areas
  (personal, home, automotive, buildings, factories, cities, ...)
- Current limitations to wider adoption of smart cameras

Panelists (expertise)

Christian Fabre
- embedded software
- low power computing

Charles Attwood
- real-time surveillance
- machine vision

Piotr Szczuko
- audio/video processing
- machine vision, AI

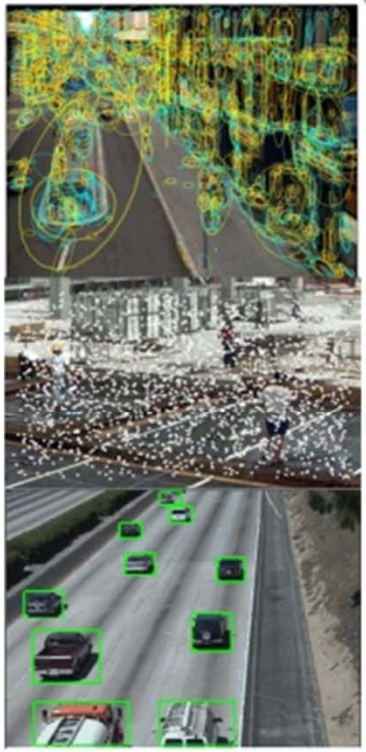
Toygar Akgun
- mapping for GPU, multi-core CPU
- parallel image/video processing

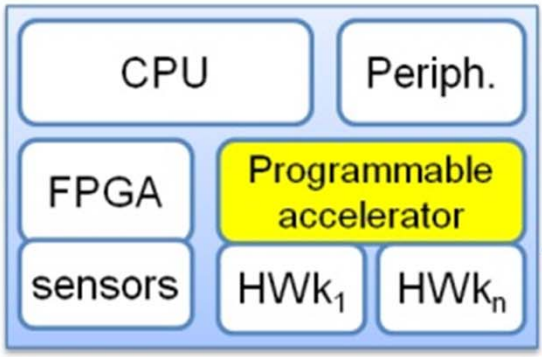
Acknowledgements
COPCAMS: COgnitive & Perceptive CAMeraS
ARTEMIS project - grant agreement No. 332913
http://www.copcams.eu

Christian Fabre

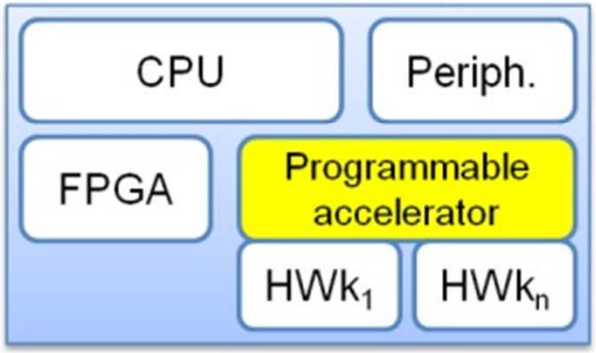
CEA LETI, Grenoble, France

Concept

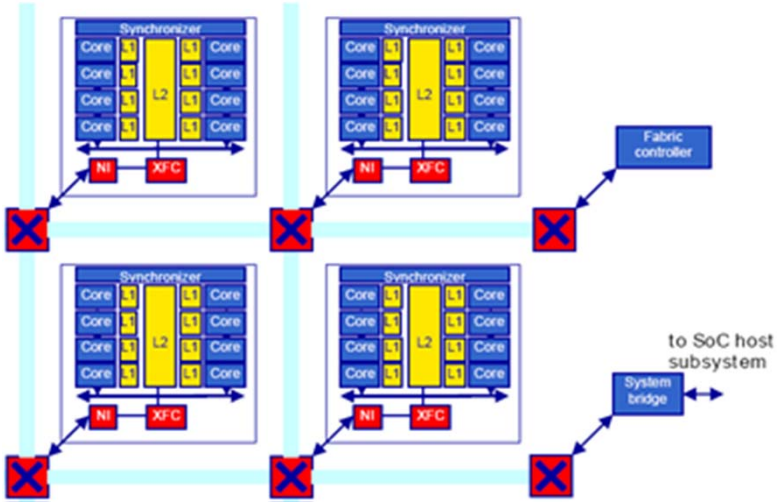




Cameras



Aggregator



Conflicting driving factors

- Low power is mandatory
  - large-scale distributed system of small (mostly autonomous) devices
  - communication-intensive systems

- Flexibility vs. efficiency
  - distributed systems
  - the more computing power, the better
  - HW blocks: computationally powerful & energy efficient but not flexible
  - SW: very flexible but less efficient w.r.t. energy & computing power
  - embedded resources are always scarce (e.g. on-chip memory)
  - networking of all kind is a must
  - bandwidth must be used efficiently
  - security is a must
  - systems are long lived: any upgrade path must be incremental

Industrial and societal aspects

- Industrial constraints
  - semiconductor costs is in *o(silicon size in* mm^2*)*
  - the more advanced the semiconductor technology is, the more volume is needed to break even
  - most SoC markets are a niche compared to the SoC market for mobile (smartphones and tablets)

- Privacy
  - data gathering?
  - reversible anonymisation
  - …

Issues at platform level

- Embedded means limited power budget
  - desktop GPGPU: 150 W for 350 Gflop/sec
  - embedded multicore: 5 W for 80 Gflop/sec
- What is the right HW/SW mix?
  - purely embedded software solution unlikely
  - find out quickly what could/should be implemented in dedicated HW

SW generic issues

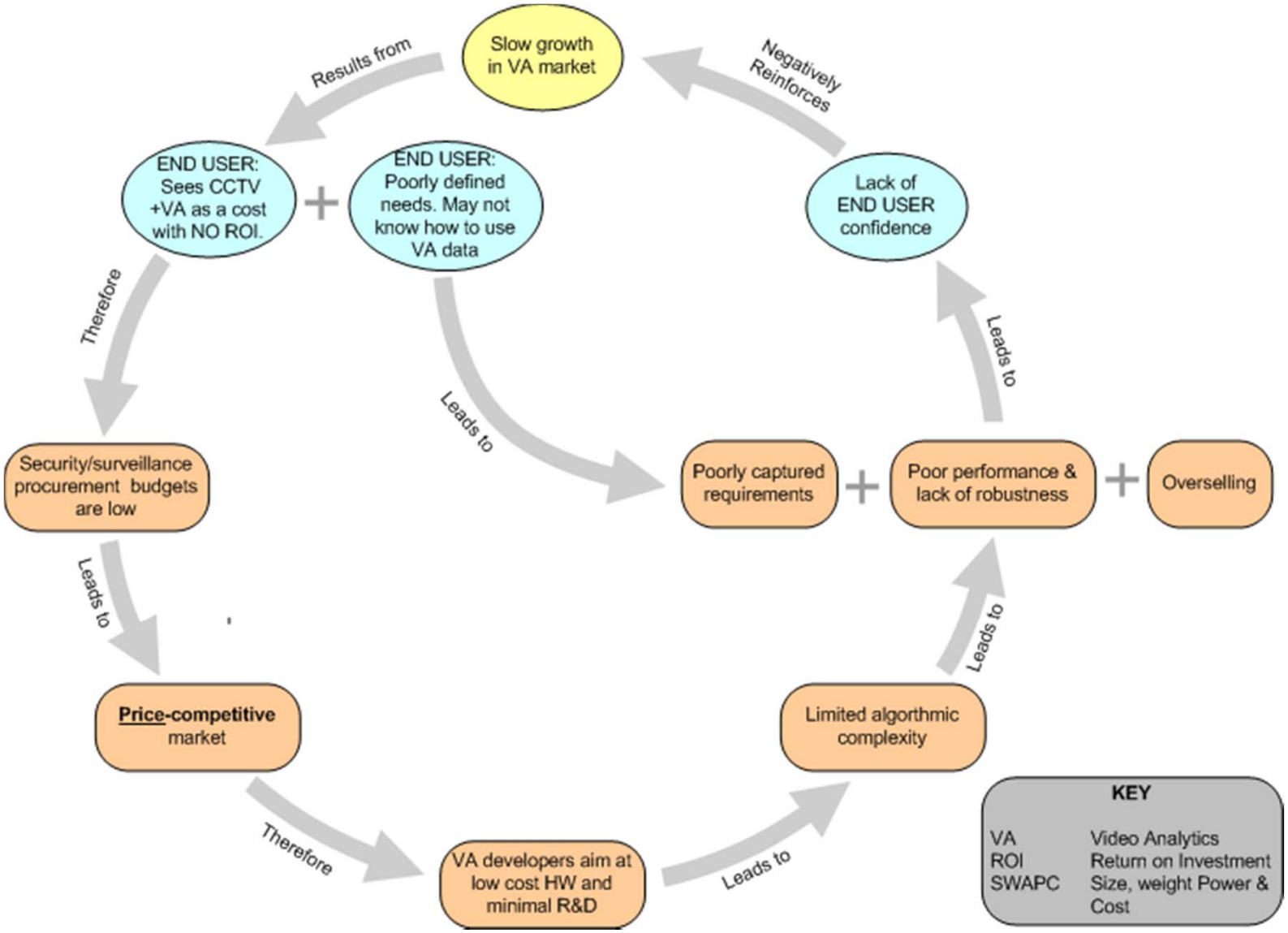
- Parallel programming
  - not easy!
  - emerging technologies:
    MCAPI, OpenACC, OpenCL, OpenCV, OpenNI, OpenMP, …

- Deployment
  - complicated!
  - distributed (heterogeneous) systems
  - integrity of the system must be preserved
  - access to the system must me checked
  - update shall be live and incremental
  - state of art in software deployment moving
    from middleware stacks to “app stores”

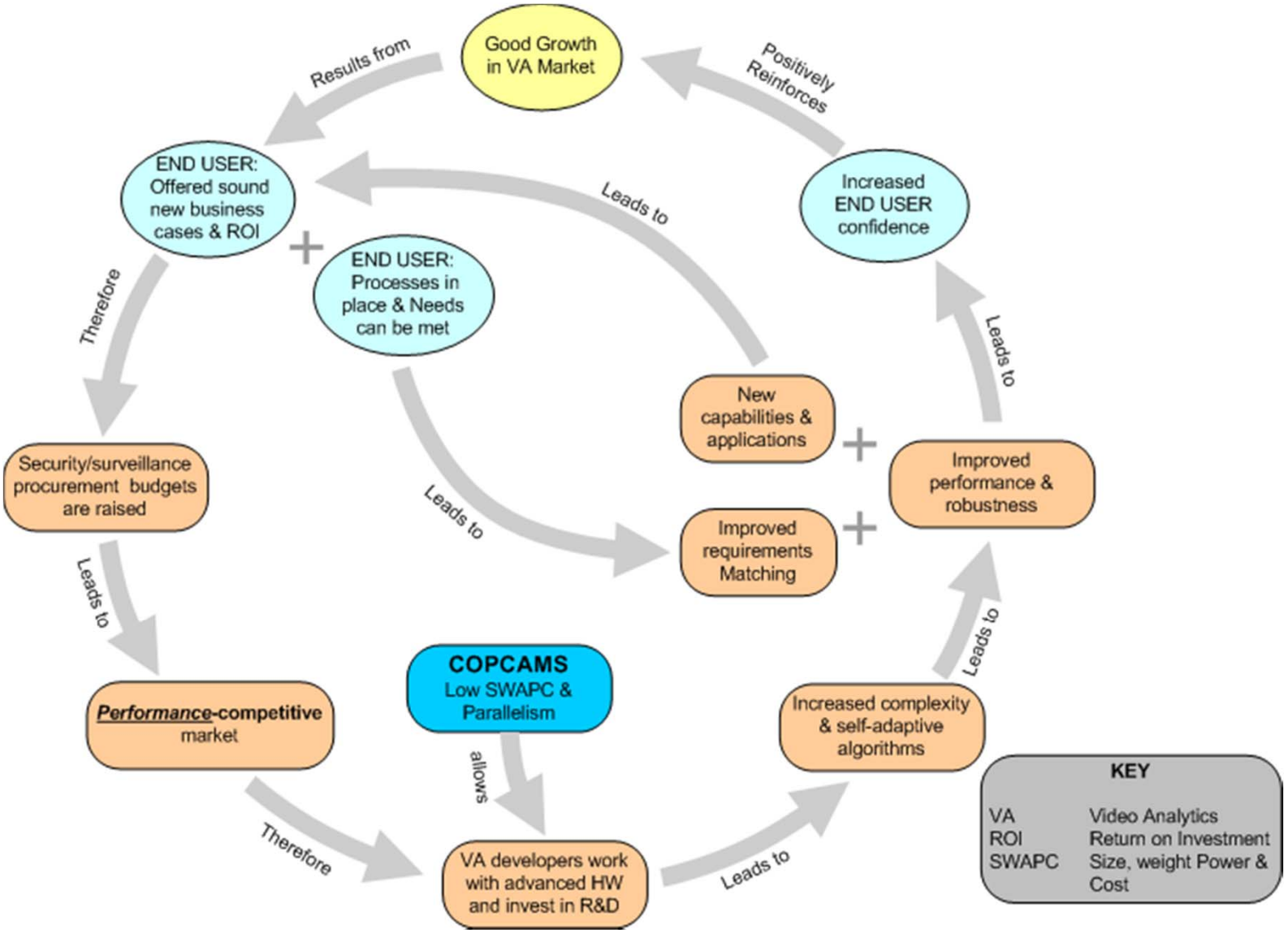
Charles Attwood

THALES-UK, Reading, United Kingdom

Perceived market failure



Market success



Piotr Szczuko

GUT, Gdansk, Poland

Challenges and limitations

- Privacy
  - real-time anonymisation
  - the original is encrypted, stored in a watermark/metadata
    - face/head/license plates localization and blurring
    - live stream anonymized, only authorized personnel can access the original *post-factum*
- Deployment
  - camera calibration, algorithm learning – *should be automated*
  - bandwidth management in surveillance applications
    - sending all metadata and low-res video (for general awareness)
    - easily navigable hi-res video on demand, multi-camera synchronized replay of archived data, etc.
  - embedded/distributed processing = scalability

Future opportunities

- Image quality
  - multi-camera, HDR, wide angle, hi-res, high fps, …
  - *how to process such a large amount of data?*
- Integration with other modalities
  - sound: identification, localization for automatic PTZ control
  - RFID: detection, identification
  - thermal / night vision
  - *how to perform data fusion?*
- Semi-intelligent, autonomous, adaptive
  - *mimic human visual/sensory system?*

Smart = Detect + Understand + Adapt + Act

Toygar Akgün

ASELSAN, Ankara, Turkey

Structural trends & issues

- Better sensors
  - more data to process
  - higher memory BW
- Higher than ever compute intensity and memory bandwidth requirements
  - must exploit data and task level parallelism
- Central *vs.* distributed
  - more processing at the end-nodes, smarter cameras
  - less visual data to be viewed/processed by operators

Structural trends & issues

- Discrete components *vs.* SoC
  - faster design cycles, better power efficiency
  - more dependence on the platform provider
- Multi-core
  - GPUs on embedded platforms
  - massive multi-threading on embedded (OpenCL/CUDA)
  - impedance from HW engineers / board designers
- Lack of multi-thread SW certifications (DO-178B)

Algorithm/application trends & issues

- Unified memory space (ARM/CPU – GPU – FPGA)
  - holy grail for OpenCL/CUDA
- Programmability – Performance – Power efficiency
  - use of fixed function HW blocks
    - *how to choose?*
  - special purpose instructions
    - *how to choose?*
    - *how to incorporate into compiler?*
  - OpenCL/CUDA drivers on embedded platforms
  - impedance from algorithm designers
- Reliable datasets & evaluation metrics for multicam systems

Are smart cameras ready for prime time?

Q&A session

Moderator

Andrea Cavallaro, *Queen Mary University of London*, UK

Participants

Christian Fabre, *CEA LETI*, France

Charles Attwood, *Thales Research and Technology*, UK

Piotr Szczuko, *Gdańsk University of Technology*, Poland

Toygar Akgün, *Aselsan*, Turkey

Biographies

Andrea Cavallaro

www.eecs.qmul.ac.uk/~andrea
andrea.cavallaro@eecs.qmul.ac.uk

Andrea Cavallaro is Professor of Multimedia Signal Processing and Director of the Centre for Intelligent Sensing at Queen Mary University of London, UK. He received his Ph.D. in Electrical Engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, in 2002 and the Laurea (Summa cum Laude) in Electrical Engineering from the University of Trieste in 1996. He was a Research Fellow with British Telecommunications (BT) in 2004/2005 and was awarded the Royal Academy of Engineering teaching Prize in 2007; three student paper awards on target tracking and perceptually sensitive coding at IEEE ICASSP in 2005, 2007 and 2009; and the best paper award at IEEE AVSS 2009. Prof. Cavallaro is Area Editor for the IEEE Signal Processing Magazine and Associate Editor for the IEEE Transactions on Image Processing. He is an elected member of the IEEE Signal Processing Society, Image, Video, and Multidimensional Signal Processing Technical Committee, and chair of its Awards committee. He served as an elected member of the IEEE Signal Processing Society, Multimedia Signal Processing Technical Committee, as Associate Editor for the IEEE Transactions on Multimedia and the IEEE Transactions on Signal Processing, and as Guest Editor for seven international journals. He was General Chair for IEEE/ACM ICDSC 2009, BMVC 2009, M2SFA2 2008, SSPE 2007, and IEEE AVSS 2007. Prof. Cavallaro was Technical Program chair of IEEE AVSS 2011, the European Signal Processing Conference (EUSIPCO 2008) and of WIAMIS 2010. He has published more than 130 journal and conference papers, one monograph on Video tracking (2011, Wiley) and three edited books: Multi-camera networks (2009, Elsevier); Analysis, retrieval and delivery of multimedia content (2012, Springer); and Intelligent multimedia surveillance (2013, Springer – to appear).

Christian Fabre

www.leti.fr
christian.fabre1@cea.fr

Christian Fabre received the Engineering Degree (MSc) from ENSIMAG (Grenoble, Fr) with a major in "Computer architecture and parallelism" 1990, while working on a Transputer routing kernel. He joined the OPEN SOFTWARE FOUNDATION Research Institute in 1993 to work on ANDF, an intermediate language. Later he worked on Java Virtual Machines, Java compilation and embedded software components. He shared the “Best Embedded Java Product” with other members of the OSF-RI TurboJava team at JavaOne in 2000. After the acquisition of the OSF-RI by Groupe SILICOMP (now ORANGE Business Services) he was part of the corporate SILICOMP R&D team. Since 2004 he has transitioned from compilation and software development to system development by adopting the MDA/MDE approach for top-down co-design of mixed hardware/software systems. He joined CEA LETI/DACLE in 2009. He has been involved in various collaborative projects, such as OMI-GLUE (1992-1995), Pastoral (2000-2012), and Expresso (French RNTL Project, 2001-2003). He is currently the coordinator of PRO3D (2012-2012, http://pro3d.eu) for DACLE/LIALP, a FP7 project using STHORM as the target platform.

Charles Attwood

www.thalesgroup.com/uk
charles.attwood@uk.thalesgroup.com

Charles Attwood is a Chief Technical Consultant at TRT-UK where he leads the Vision Lab team. After graduating in experimental psychology (1981) and an MSc in Intelligent systems, he began his professional career as an IBM Research Fellow in 1986. He spent 10 years in the Computational Vision group at Reading University. He joined Thales in 1996 working on advanced studies and development of the market leading automatic number plate recognition system called Talon. He was the Technical Coordinator for the highly successful EC Framework project - ADVISOR, which developed a real-time intelligent surveillance system, trialled in Barcelona Metro. He has 10 refereed publications, and three patents in image processing. He is a past member of the Executive Committee of the British Machine Vision Association, and has sat on many UK steering committees for EPSRC projects in the area of machine vision. His current research interests are in robust visual processing with a low-computational complexity emphasis.

Piotr Szczuko

www.multimed.org
szczuko@multimed.org

Piotr Szczuko is an assistant professor in the Multimedia Systems Dpt., where he received his M.Sc degree in Sound and Vision Engineering (2002), and PhD in Telecommunication and Computer Science (2008), with his PhD thesis being awarded by the Polish Prime Minister. He is an author and co-author of 70 papers in journals and international conferences. His research activities are mainly audio and video processing for automatic detection of threats in city agglomerations. He is also interested in: machine vision, artificial intelligence and automatic inference methods, and classification and perception of sound and image.

Toygar Akgun

www.aselsan.com.tr
takgun@aselsan.com.tr

Toygar Akgun received the B.S. degree in Electrical Engineering from Bilkent University, Ankara, Turkey, in 2001, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the Georgia Institute of Technology, Atlanta, GA in 2004 and 2007, respectively. He was with NVIDIA's Windows Video Group between 2008 and 2011. He joined ASELSAN in 2012 to work on video and signal-based security and surveillance. His research interests include digital image/video processing, algorithm development, analysis and mapping for GPUs, task and data parallel processing of digital image/video signals, statistical signal processing and modelling.