C O P C A M S

Cognitive & Perceptive Cameras

Artemis-JU GA n°332913

D5.3 – Outdoor & Indoor Building Surveillance Applications Specification

WP5 –Applications & Field Tests

Authors:

Alberto Isasi – TCNL, Christian Fabre -CEA

Date: 2016-01-05 at 16:24

Version: v2

Status: Public

CEA ref. number: LETI/DACLE/15-0904

http://copcams.eu

Document History

Date	Author	Modification
2015-03-18	Julie Foucault (CEA)	Template creation
2015-04-03	Bogdan Filipič (JSI)	Table of contents
2015-05-15	Alberto Isasi (TECN)	Traffic surveillance content
2015-09-08	Piotr Szczuko (GUT)	Building surveillance content
2015-10-07	Charles I. Attwood (TRT-UK)	Robust Motion Detector content
2015-10-20	Juan L. Sanchez (CCTL)	Smart Building content
2015-10-23	Darryl Hond (TRT-UK)	Motion Detection content
2016-01-05	Julie Foucault(CEA)	Added chrono number

Contributors

Alberto Isasi, Adrian Galdran, Antonio Cruz (TECN)

Piotr Szczuko, Andrzej Czyzewski (GUT)

Charles I. Attwood, Darryl Hond (TRT-UK)

A. Montalban; J.L. de Amaya; J. Llinares; F. Sacristán; A. Alcaide; E. Saltalamachia; Juan L. Sanchez (CCTL)

COPCAMS Partners

1	CEA	Commissariat à l'énergie atomique et aux énergies alternatives
~~2~~	~~ST-FR~~	~~STMicroelectronics (Grenoble 2) SAS~~
~~3~~	TCS	THALES Communications & Security SA
4	TRT-FR	THALES Research & Technology France (repr. THALES SA)
5	INRIA	Institut national de recherche en informatique et automatique
~~6~~	~~IIMS~~	~~Information & Image Management Systems~~
7	CTTC	Centre Tecnolòlogic de Telecomunicacions de Catalunya
8	CCTL	Concatel
9	IQU	Iquadrat Informatica S.L.
10	TECN	Tecnalia Research & Innovation
11	TED	Tedesys Global S.L.
12	UC	Universidad de Cantabria
13	GUT	Politechnika Gdańska
~~14~~	~~BS~~	~~BS Spółka z ograniczoną odpowiedzialnością Sp. k.~~
15	JSI	Institut ``Jožef Stefan"
16	DTU	Danmarks Tekniske Universitet / IMM
~~17~~	~~ASLV~~	~~Application Solutions (Electronics and Vision) Ltd~~
18	TRT-UK	THALES Research & Technology (UK) Ltd
19	QMUL	Queen Mary University of London
20	ASEL	ASELSAN Electronics Industry
21	KTOR	Kolektor Group d.o.o.
22	SOG	Sogilis
23	SQST	Squadron system
24	WLNS	Wavelens

Table of Contents

1 Introduction

This document is deliverable D5.3. Its goal is to describe the different scenarios and use cases for the WP3 developed algorithms regarding to Indoor and Outdoor Surveillance (buildings and traffic domains).

To achieve this objective the different experiments carried out in the labs will be described and it will also be pointed out how these experiments will be carried out to real environments. It will be explained the different architectures selected for each use case and the methodology for its correct implementation and validation.

Regarding to the systems validation, it will be described what was the state of the art at the beginning of the project and how it has been improved with the different development carried out in the project.

This document finishes summarizing the current status of the different demonstrators and the conclusions extracted from the results of the validation tests.

2 Demonstrator task

In this point, the different demonstrators developed by each partner will be addressed. As in this task there are different environments and challenges to be faced (indoor/outdoor surveillance, video + audio data fusion, day/night motion detection, image enhancement…) each demonstrator will include a brief description of itself and what are the improvements to achieve.

2.1 Building Surveillance Scenario

Current generation of surveillance used for monitoring of safety and for access control in buildings and their vicinity is based on applying large number of cameras, allowing covering all areas of interest, to continuously track objects in operator's focus. It is common to use additional, yet rudimentary sensors such as infrared motion detectors. Such a standard setup is operated by means of concurrent observation of as many video streams as practically feasible on LCD screens. Thus, the final accuracy is hampered by the number of cameras, number of operators and their attention timespan, size of covered area, video streaming quality, latency (delays), and possible other factors.

The developed prototype of "cognitive and perceptive camera" aims at reducing aforementioned negative influences, by:

- Limiting the necessity of direct observation of the video feed by the user,

- Reducing number of cameras required for accurate tracking by employing statistics and object re-identification methods,
- Decreasing amount of streamed data and detection delays by processing the video at the edge (processing unit inside the camera casing or near the camera with a wired connection),
- Increasing number of possible detected events by introducing new types of sensors: acoustic and RFID-based ones.

2.2 Traffic Surveillance Scenario

TECNALIA use case has been defined for traffic surveillance and managing. There are two main objectives in this demonstrator: Detect & Classify vehicles (trucks, cars, motorbikes...) under any environmental situation (day/night, sunny/cloudy day, rain/snow/fog…) and Detect a set of incidents in tunnels automatically (AIDs). The location for the system deployment will be placed in Txoriherri motorway with an outdoor setup. For tunnels there will be another different setup.

The main task for the image processing algorithms can be summarized in:

- Image enhancement for image quality normalization (contrast stretching, colour normalization, dark channel dehazing…)
- Movement objects' detection. Compute the dynamics (velocity and acceleration) for each moving object.
- Detected objects' classification. Classify each moving object into a set of vehicle classes.

For Automatic Incident Detection in tunnels the goal is to detect the following situations:

- Detection of stopped vehicles
- Detection of vehicles driving in wrong way
- Detection of people walking along the tunnel

2.3 Building Surveillance Scenario (RMD)

Two completed TRT-UK COPCAMS tasks are referred to in this deliverable. Both of these involved activities related to an established Robust Motion Detector (RMD) and both are described at greater length in COPCAMS deliverable D3.5 [2]. The first task entailed the production of a modified version of the RMD. The second required the RMD to be ported onto a development board. Both of these tasks have led to progress which can be demonstrated in some manner. The level of performance attained by the modified RMD is demonstrated by applying it to two databases. The manner in which the ported RMD can be demonstrated is presented in section 2.3.1.

2.3.1 *Demonstrating the Ported RMD*

Three approaches can be taken when demonstrating the RMD on the Embedded Platform.

1) Processing video data on the platform.

2) Using a USB webcam for data, running the algorithms in real time and outputting the results to the on-board display.

3) Offline processing can be used to demonstrate the forensic potential of the RMD which ensues from its ability to abridge long video clips. For example, a 24 hours clip can be condensed to a few minutes of motion-filled footage.

Note: TRT-UK intend to demonstrate the ported RMD at the COPCAMS Gdansk review, November 2015. The presentation will draw on the second mode of operation discussed above, processing data captured live by a webcam.

2.4 Cognitive and Perceptive Vision Systems for Smart Building Scenario

Smart Facility and Building Management generally involves a number of disciplines and services. The most general description to identify the market segment is to understand Smart F&BM as the integrated management process that considers people, process and place in organisational context, being focused in the design and improvement of intelligent buildings (IB) and the coordination and optimization of several domains: facilities, life security, physical security and information technology.

One of the key trends in Smart Facilities & Building Management is to provide solutions that can take remedial actions automatically, providing a coordinated response in the “foundational systems” such as security, electrical and lighting distribution or HVAC (heating, ventilation, and air conditioning). For instance, the heating, cooling and ventilation represents for 30% of energy usage and for 50% of the electricity. Currently, most modern buildings still condition rooms assuming maximum occupancy rather than actual usage. As a result, rooms are often over-conditioned. Temperature and CO2 levels are two main conditioning factors to consider for HVAC control strategies. Temperature only requires a binary indication if a room is occupied, which could be implemented using a Passive Infrared Sensor (PIR). However, CO2 ventilation rates are a function of the number of occupants and cannot be effectively controlled via PIR. CO2 sensors are problematic for ventilation control since they may require a significant amount of time before detecting CO2 buildup. Thus, an HVAC control strategy must utilize an occupancy monitoring system capable of detecting the number of occupants in real time.

In this sense, the video surveillance system, that is a mainstay of building security, could serve also for this purpose. The analysis of digital images addresses aspects of physical security but may go way beyond that to provide data and information for building life safety, energy management and overall building performance.

In this context, the Cognitive & Perceptive Video Systems (CPVS) enabled by COPCAMS would represent a significant step towards wider adoption of embedded vision systems within the smart facilities & smart building management domain. COPCAMS could provide an HVAC control strategy based on occupancy prediction and real time occupancy monitoring, as an added value feature, beyond the current state of the art in smart surveillance systems.

The motivation in this use case is to test and iteratively improve the approach (together with the use case in T5.1: Cognitive and Perceptive Cameras Systems for Smart Building Management), in order to identify a minimum viable service (MVS) that can be provided to different clients as a comprehensive solution within the Smart Facilities and Smart Buildings Management domain, taking into account performance, cost, efficiency and deployment requirements.

Specifically, this use case is based on the potential use of a CPVS to provide different functions/profiles depending of different situations. That is, to explore the potential of COPCAMS approach, -with embedded and powerful vision systems- to sense the surrounding environment, and react to changes. In this case, the field test aims to explore the possibility of a COPCAMS system that is initially working as part of an Occupancy Based Demand Response HVAC, ("HVAC mode") to change to "surveillance mode" to detect intrusions in the specific indoor zone.

Figure 1 – Different room occupancy types.

This new approach would provide advanced features in an emerging market, where and improved performance and reduced energy consumption will facilitate the use of embedded cameras not only as simple sensors, but as a distributed cognitive system, going beyond smart surveillance.

3 Related lab experiments

This point will address the different experiments to carry out to test and validate each demonstrator. It includes the experiments developed both in PC as in the selected embedded platforms.

3.1 Building Surveillance Scenario

The event detection and accuracy assessment will be made in accordance with the following scenario.

In the test area a typical movement of people will be staged - individuals walking along the corridor in both directions, varying movement speed. In the field of view of camera “A” a high security area will be arranged, including a table, cabinet, chair, coat hanger, desk, cupboard, and housing a valuable object of interest (equipment), tagged with the RFID tag. The event of stealing the object of interest will be simulated, by playing sound of breaking glass from a nearby speaker and taking the object away and leaving the scene through parts observed by cameras A, B, C and P.

Camera A equipped with audio sensor and RFID transceiver will be used to detect and localize sound of breaking glass and RFID tag displacement.

All objects visible in cameras A, B, and C will be tracked, and the object within the area of the detected event will be associated with the theft and marked as a culprit. Among detected and tracked objects the culprit will be only object of interest and PTZ camera P will be automatically steered to follow it.

General overview of the scenario is presented in the following image:

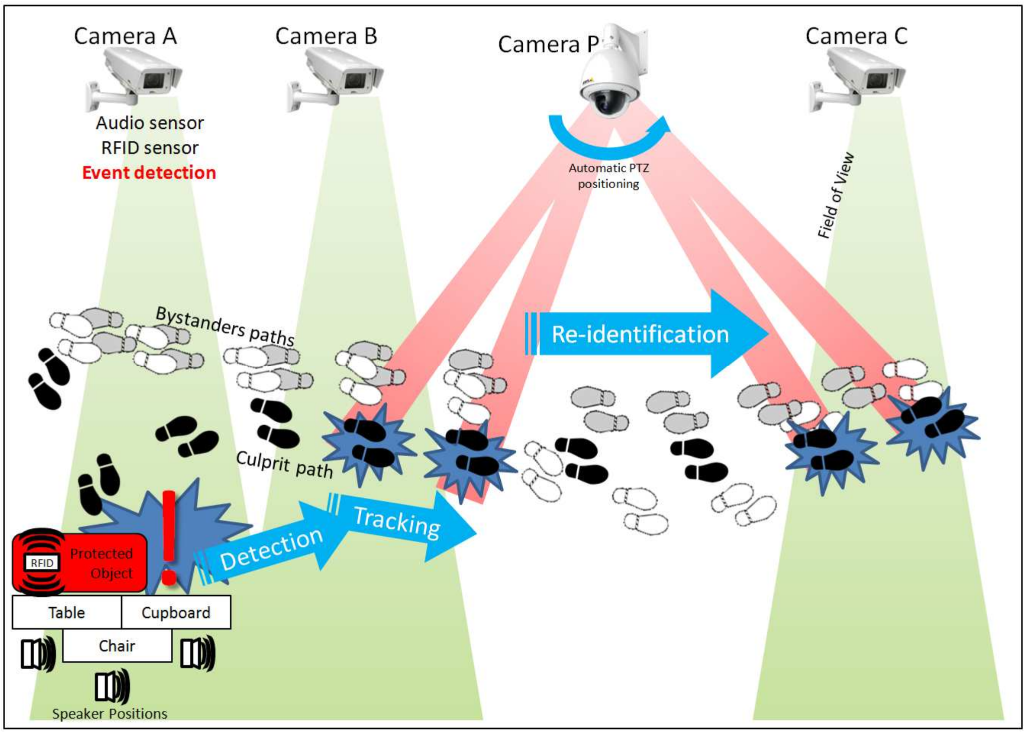

Figure 2 – Testing scenario scheme.

The test scenario will be performed and evaluated 60 times with spontaneously varying behaviour of bystanders (Table 1). The recording will last for 3 hours, in two 1.5h sessions including 30 performances. Each performed action should take 1 minute at most, followed by 2 minutes of typical behaviour, such as idling and casual walking of bystanders. The controlled conditions will be as follows:

- 25 repetitions in an evenly distributed daylight and 25 in artificial light,
- 10 repetitions of the scenario performed by 3 people, other 10 by 4 people and 10 by 5 people, all dressed casually in varying colours clothing,
- Similar 3x10 actions of 3, 4 and 5 persons but dressed officially: dark cloths, jackets, bright shirts,
- Even and odd repetitions will be performed alternatively in slow and brisk walk,

- The location of the protected object and the speaker emitting various glass breaking sounds will be changed to validate correct theft detection in wide range of places,
- Other types of sounds should be present, performed spontaneously: talks, footsteps, phone calls, door shuts, etc., typical for office environment.

No.	Light conditions	Number of people	Dress type	Walk speed	Speaker location	Emitted sound
1	Daylight	3	Casual	Slow	On the table	Glass Break A
2	Daylight	3	Casual	Brisk	On the table	Glass Break B
3	Daylight	3	Casual	Slow	On the table	Glass Break C
4	Daylight	3	Casual	Brisk	On the table	Glass Break D
5	Daylight	3	Casual	Slow	On the table	Glass Break E
6	Daylight	3	Casual	Brisk	Under the table	Glass Break A
7	Daylight	3	Casual	Slow	Under the table	Glass Break B
8	Daylight	3	Casual	Brisk	Under the table	Glass Break C
9	Daylight	3	Casual	Slow	Under the table	Glass Break D
10	Daylight	3	Casual	Brisk	Under the table	Glass Break E
11	Daylight	4	Casual	Slow	Near wall	Glass Break A
12	Daylight	4	Casual	Brisk	Near wall	Glass Break B
13	Daylight	4	Casual	Slow	Near wall	Glass Break C
...	...	...	...	...	...	...
56	Artificial	5	Official	Brisk	Behind the cabinet	Glass Break A
57	Artificial	5	Official	Slow	Behind the cabinet	Glass Break B
58	Artificial	5	Official	Brisk	Behind the cabinet	Glass Break C
59	Artificial	5	Official	Slow	Behind the cabinet	Glass Break D
60	Artificial	5	Official	Brisk	Behind the cabinet	Glass Break E

Table 1 - Test cases description

3.2 Traffic Surveillance Scenario (TCNL)

For testing the algorithms developed for traffic surveillance, a set of videos acquired by the acquisition system deployed in the gantry sign will be used. Those videos include different environmental situations such as sunny day, cloudy day, rainy day, at day and night… so we can get the response of the algorithms for each situation.

The first experiments have been done in PC platform and once all the algorithms are coded in the embedded platform, they will be repeated to compare the outputs and error rates.

3.2.1 Experiments developed on PC

Using the cameras mounted in the demonstrator, a set of videos has been recorded. In those videos, there are a large number of vehicles crossing under the gantry sign structure and they have been processed using the algorithms described in documents [3] and [4] summarizing the results in the following table:

	Number of Vehicles	Detected Vehicles	Missing Vehicles	ERROR (%)
Video 1	**220**	**217**	**3**	**98,6%**
Video 2	**215**	**212**	**3**	**98,6%**
Video 3	**405**	**405**	**0**	**100%**
Video 4	**225**	**215**	**10**	**92,89%**
On Line	**86**	**80**	**6**	**93%**

Table 2 – Vehicle detection algorithm results

There are some problems detecting vehicles under heavy rain at night conditions, being the algorithms too weak in that scenario. To improve the performance in those conditions TECN is working in a set of image enhancement algorithms to remove the effects of shines and water drops.

3.2.2 *Experiments developed on PC+GPU*

The first experiments developed in the GPU platform are related to the image enhancement algorithms. In this case, as they are using matrix operations too often and are completely paralelizable, the performance that these algorithms offer is really good. In the following images are presented the comparison between CPU and GPU performance:

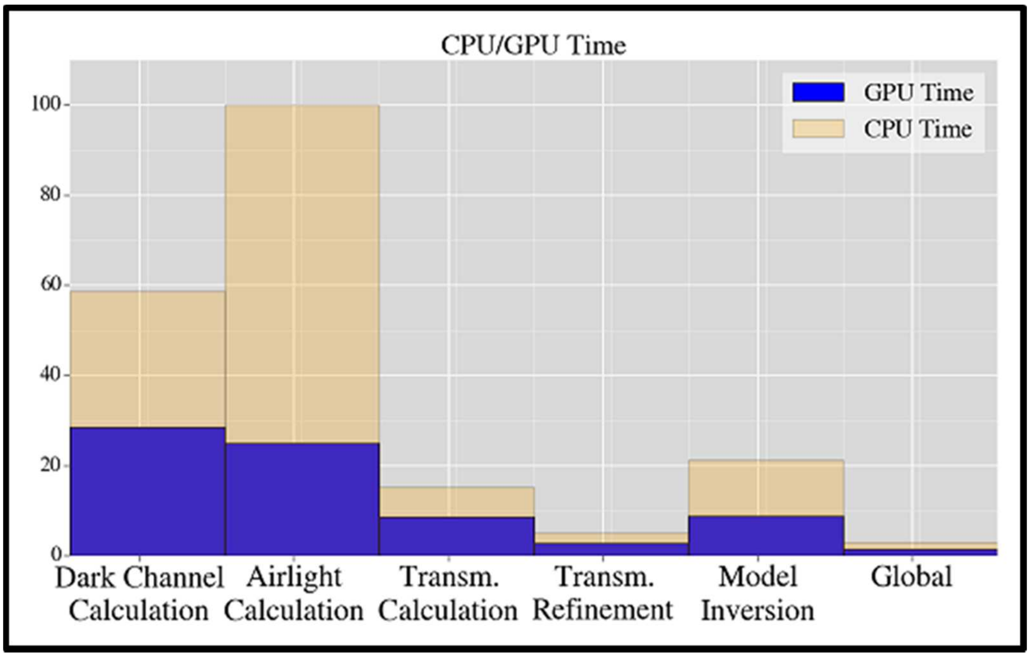

Figure 3 – Dark channel dehazing performance.

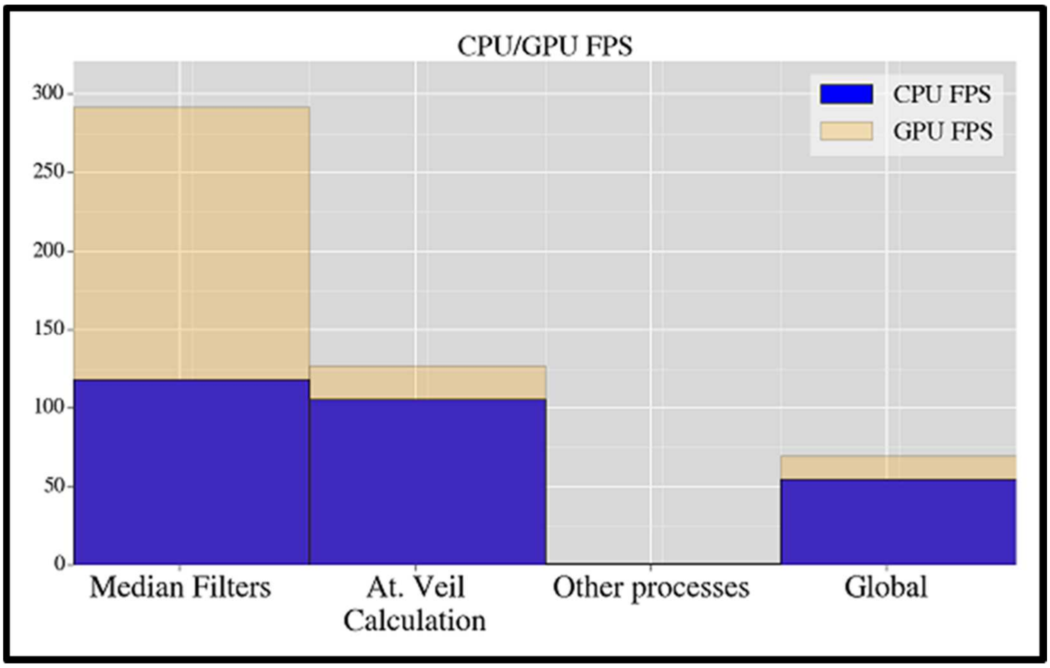

Figure 4 – Contrast Enhancement in Road Scenes.

As it is shown in the images, not all the steps are parallelizable but the ones that can be deployed to GPU offers a great performance increase.

3.3 Buildings Surveillance Scenario (RMD)

The TRT-UK Robust Motion Detector (RMD) is a patented system for motion detection. It is robust to illumination change and a degree of repetitive local motion, and requires only low computational resources. The RMD has been designed in the first instance to operate in a forensic mode, whereby live or pre-recorded video is filtered on the basis of detected motion to produce an abridged clip displaying events of interest. This clip can be reviewed in much shorter time than the original.

The goal of one TRT-UK COPCAMS sub-task was to adapt the RMD so that it could operate in an alternative, non-forensic mode. In order to modify the RMD, three candidate approaches were selected for evaluation.

In addition to this, the work to port the RMD to a development board has required an empirical assessment of the efficacy of co-processing units with regard to optimisation.

3.3.1 Experiments developed on PC

3.3.1.1 Adaptation of the RMD to act in a non-forensic mode

In order to produce an RMD which operated effectively in the desired mode, three independent approaches were conceived and evaluated.

3.3.2 Experiments developed on Embedded Platform

The RMD was ported to the development board. Optimisation followed. First, the algorithm was run on the board with simple profiling information produced to identify where the processing bottlenecks occurred. This led to identifying certain co-processor accelerator units as good candidates for removing the bottlenecks. These were then exploited in order to improve performance. The relationship between the engaged co-processor resources and various aspects of performance required empirical assessment. This took the form of recording the baseline results for the initial version, applying various combinations of the accelerator options and analysing which combinations gave the best improvements to performance. Further work would continue this approach by re-profiling and then identifying further sections to offload to accelerators.

3.4 CPVS for Smart Building Scenario

To check the feasibility of the use case, and determine the scenario structure, evaluation strategy and validation, a deep study of the state of the art has been performed, and different algorithms and approaches have been analysed. The aim was to define a set of candidate algorithms and approaches, and identify an initial configuration for the field test. This approach has taken into account the basic requirements: identify a minimum viable service (MVS) that can be provided to different clients as a comprehensive solution within the Smart Facilities and Smart Buildings Management domain, taking into account performance, cost, efficiency and deployment requirements.

In this sense, the preliminary activities performed during the lab experiments have led mainly to determine a configuration based on the occupancy detection algorithm described in “D3.5– Video Algorithms – On COPCAMS Platforms”, with a PC+GPGPU configuration. The algorithm has provided good lab results in different environments for a frame size of 352x288 and a sample rate of 12fps with a detection rate over 80% and false alarm rate of 15%.

During the field test, the need of illumination detection and background correction will be analysed, to be implemented in a commercialization phase.

Regarding human detection, different algorithms have been studied, having as reference for the implementation in COPCAMS the personnel detection algorithms described in [24] and [25].

4 Demonstrator architecture and methodology

4.1 Building Surveillance Scenario

GUT-MSD is engaged in developing technology of multimodal event detection and object tracking. Those algorithms involve processing of audio and video streams for event detection and object tracking, and RFID signals analysis for localization and tracking of objects. For the purpose of demonstrating a prototype a test environment is designed.

Prototype tests will be conducted in a selected hall of ETI faculty of GUT, employing a dedicated setup of cameras and sensors.

The hall floor area will be covered by digital cameras denoted as A, B, and C, and pan-tilt-zoom camera P. Fields of view of A and B will be overlapping to validate multi-camera tracking in case of simultaneous visibility of the object of interest in two cameras; fields of view of B and C will be located 3-5m apart, to validate tracking method enhanced by additional reasoning based on temporal and behavioural statistics. Camera P will be mounted in the middle of corridor length and positioned automatically using developed tracking algorithms, to follow object of interest.

Camera A will be equipped with an audio sensor, and a RFID sensor. The main task of this unit is to detect important, security-related events, defined in the next subsection. It will be used for detection and object localization, and RFID tags localization and tracking, resulting in user notifications and automatic camera positioning, and movement tracking involving automatic camera positioning.

Cameras B and C will be equipped with a RFID sensor, and will be used for object's visual reidentification and tracking, and RFID tags localization and tracking, resulting in user notifications and automatic camera positioning.

Digital IP cameras by AXIS will be used. The audio sensor and RFID sensors prepared by the GUT-MSD team will be employed.

4.2 Traffic Surveillance Scenario

This scenario has two main use cases:

- Free flow vehicle detection and classification
- Automatic Incident Detection (AID) in tunnels

The selected location for the system deployment is sited in Txoriherri motorway, near Tecnalia's facilities. For AID algorithms, TECNALIA has been speaking with Bizkaia Regional Government in order to capture some videos from Bizkaia's tunnels. The have provided some videos from different tunnels for testing porpoises. The same had been done with motorway deployment: To be able to compare the output from different platforms, a set of videos in different environmental situations have been acquired.

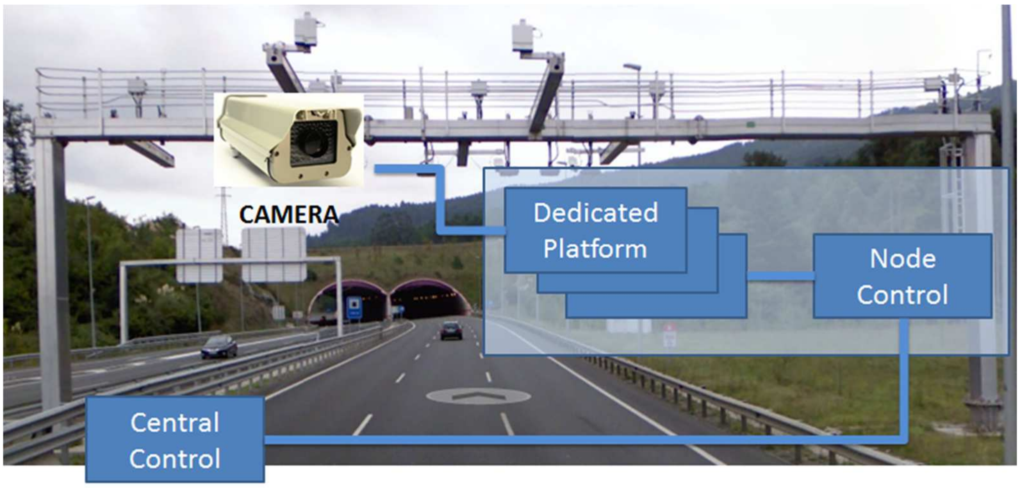

Figure 5 – Traffic surveillance deployment.

The methodology to follow during the experiments will be as follows:

- For each platform a set of acquired videos will be processed. These videos will contain different environmental situations and incidents.
- The measures to be evaluated will be:
  - Success rate
  - Error rate
  - False positives rate
  - False negatives rate
  - Frames per second (max, min, mean)
  - Power consumption

Each measure will be pondered in order to its relevance and impact in the application and the different algorithms will be evaluated.

4.3 Building Surveillance Scenario (RMD)

The two different platforms used at TRT-UK have different architectures and methodologies.

4.3.1 Experiments developed on PC

The main development environment was Visual Studio/ C++. Some experimental code was written in Matlab.

4.3.2 Experiments developed on Embedded Platform

The experiments on the development board used Yocto as the board support package, gcc and makefiles to build the application, and gdb and valgrind for debugging and profiling on target, respectively.

4.4 CPVS for Smart Building Scenario

The field test scenario is initially simulated in CCTL facilities, and the setup of the system is as follows: A COPCAMS platform (initially a PC+GPGPU, but extensible to STHORM platform) with a single camera will be placed to monitor an indoor working area, initially working as part of an Occupancy Based Demand Response HVAC ("HVAC mode"). This process will be based on the algorithms implemented in WP3, having as reference the proposal described in [26]. On a particular moment (triggered by the end of working time or by a specific simulated alarm that will be captured by the system), the system will be required to change to surveillance mode, and detect human intrusions in that working zone. The results will be registered, in order to be sent to a main station that could feed a smart HVAC system for occupancy prediction, a business intelligence unit, a dashboard or a decision-making system.

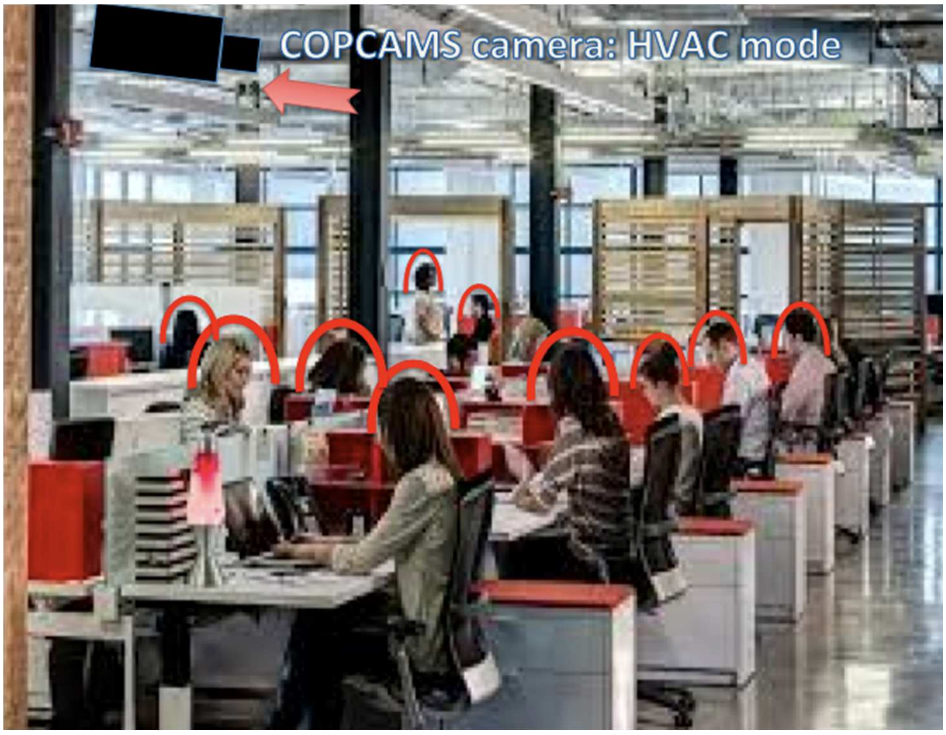

Figure 6 – Occupancy detection for Smart HVAC.

The described scenario will take place in an indoor controlled area, with stable illumination. We assume that the system is correctly installed to cover the monitoring area and the cameras are calibrated.

Under these conditions, the following metrics will be measured for the performance evaluation.

- **Detection Rate:** For both modes, the detection rate will be measured.
- **False Alarm Rate:** For both modes, the false alarm rate will be measured.
- **Frame rate:** The minimum and maximum frame rate required to provide a valid service.

The evaluation strategy will follow an iterative process, where the algorithms and the overall COPCAMS performance (including power consumption) will be analysed in different target platforms.

5 Validation

5.1 State of the art at project start

5.1.1 Embedded computer vision for traffic solutions

Among digital signal processing markets, embedded vision is considered one top-tier, fast growing area [10, 11]. Embedded vision refers to the deployment of visual capabilities to embedded systems for better understanding of 2D/3D visual scenes [10]. It covers a variety of rapidly growing markets and applications.

Emerging high-volume embedded vision markets include automotive safety, surveillance, and gaming. Similar functionality may be useful in a variety of systems targeting different markets.

While vision computing is already a strong research focus, embedded deployment of vision algorithms is still in a fairly early stage. The embedded realization of vision algorithms is notoriously difficult. Vision computing increases the demand for extremely high performance, coupled with very low power and desire for low cost.

For video surveillance, the estimation shows more than 3-fold growth from $11.5 billion in 2008 to $37.5 billion in 2015 [12]. An even stronger growth is predicted for the ADAS (Advanced Driver Assistance Systems) market with a 14-fold growth over eight years ($16.6 billion in 2012 to $261 billion in 2020).

In intelligent transportation systems (ITS), computer vision technology is broadly applied to either 1) detect objects and events that may represent safety risks to drivers, or 2) improve the efficiency of road networks.

In case 1) applications are focused to self-navigated for autonomous Advanced Driver Assistance Systems, ADAS and in case 2) are focused to traffic and road management

The number of vehicles on roads keeps increasing continuously, making the management of traffic flow, especially in big cities more and more challenging. One of the key enablers for having smooth traffic flows and better mobility is to rely on real-time traffic monitoring systems. These systems allow road operators to implement intelligent traffic management strategies such as the dynamic adjustment of timing and phasing of traffic lights and the adaptive road congestion charging.

ADAS systems are commercially available today and next-generation systems promise great technological advancements. Consumer demand is high and the road to widespread adoption is being paved. In fact, Visiongain forecasts that the global ADAS market will experience double-digit growth between 2014 and 2024, from a baseline estimate of $18.2 billion [13]. Some hurdles to bringing next-generation ADAS systems to mainstream vehicles remain, but given the remarkable strides made over the last few years and the burgeoning ADAS ecosystem, the industry can and will overcome these challenges.

The systems will use new, purpose-built ADAS SoCs that provide support for the multiple sensor technologies. The wide availability of these deep-submicron SoCs, built by companies such as Freescale, Intel, NVIDIA, Renesas, and Texas Instruments, will spur system designers to replace discrete, single-function ADAS systems implemented on low-end microcontrollers with smarter, consolidated multi-sensor systems that can handle high data throughput in real time.

These SoCs often incorporate multiple CPU cores, digital signal processors, general-purpose graphics processing units (GPGPUs), or vision acceleration engines, along with multiple camera inputs and display outputs. The SoCs will offer automakers increased flexibility and control over how their ADAS systems are designed, enabling commercially attractive and differentiated solutions.

Traffic and road management applications are focused to improve traffic flow and safety; these applications have to perform tasks during daylight, low sun, darkness and in extreme climatic conditions, so the complexity is high. Include functionalities:

- ITS (intelligent transportation systems), Traffic Analysis, Vehicle Count, Speed Checks
- Traffic patterns, wait times and lengths
- Automatic traffic reports and congestions
- Vehicle Parking Violations
- Parking Lot Management
- Stopped Car Detection
- Automated License Plate Detection
- Automated Vehicle Detection

Capability drives cost. This has been a concern within the ITS market, which has traditionally been used to lower-end offerings more like those used in the security and general surveillance sectors. This resulted in machine vision's use in traffic management being restricted to high-end, such as tolling and enforcement applications, however this is now changing and it is the low-end applications of machine vision technology which are the current pre-eminent trend.

Machine vision industry is taking traffic installations seriously, this is evident by the amount of hardware and software products tailor-made for ITS applications that are now available on the market [14].

The development of the computer vision algorithm only represents one part of all the product's design cycle. One of the hardest tasks is to validate the whole system with the wide variety of driving scenarios. Usually, more time is spent in testing and validating a system than developing the algorithm. In order to validate a computer vision-based solution, thousands of hours of video are necessary. Suppliers spend a lot of time gathering videos that cover all the possible scenarios, which implies recording in different illumination, weather and road conditions. Once the video compilation is finished, it is necessary to annotate all the objects and events that have been recorded and need to be detected by the algorithm. Traditionally, this video annotation has been done manually, having dedicated workers labelling each frame. However, as the size of the video databases grow, this solution becomes non-feasible.

5.1.2 Embedded computer vision solution for ITS market

Applications more mature, with commercial solutions are traffic enforcement, tolls and vehicle recognition and AIDs (Automatic incident detection system) for tunnel safety. Most of them involve ANPR (Automatic Number Plate Recognition) but in AID system also fire and other incident detection are implemented.

In AIDs solutions, key players as Citilog [16] has PC based solution and FLIR [17] has a camera with embedded SW.

Regarding ANPR, also coexist PC based solutions and embedded technology in which information is entirely processed on board the smart camera.

Some of the key players in the global automatic number plate recognition marketplace include PIPS Technology Ltd., NDI Recognition Systems Ltd., Genetec Inc., Tattile Srl and Bosch Security Systems. The key strategy used by players to sustain in the ANPR market is to develop ANPR solutions with better reading capabilities and integrated ANPR systems. The market is highly fragmented and thus, the competition within a particular market segment such as ANPR engine manufacturing, ANPR component manufacturers and system integrators is high. In the coming few years, manufacturers are expected to expand their operations in emerging economies such as China and India as the ANPR adoption in these countries is expected to swell over the forecast period [18].

Most of the embedded solutions implemented in the cameras are DSP and FPGA based with clear and limited functionalities and have not evolved in the few past years.

Although, another technological approaches like connected car technology is evolved and seems a crucial part of the future of transport and road safety. New developments are appearing for new real-time traffic monitoring based on emerging vehicular communication systems, V2V (vehicle-to-vehicle) and V2X (vehicle-to-infrastructure) [19]. New systems enable traffic monitoring with higher reliability, accuracy, and granularity. Companies, as DENSO, leader in the latest V2V (Vehicle-to-vehicle) and V2X technology allows cars to “talk” or communicate with other surrounding vehicles and traffic signals through wireless communication and data exchange in the combinations.

Also, due to embedded vision for ADAS is at the moment a very active area of research, a lot of efforts and new products are appearing and in the coming years, new innovative solutions are expected to appear together with smarter tools to validate them. Most of these solutions involve computer vision technologies pattern recognition, image processing, real-time signal processing.

We can assume that new implementations and progress in this field will be translated to traffic and road applications.

In the case of ADAS solutions, there is not a clear winner among the different candidates for being the referent embedded platform for vision-based ADAS but it is an increasing trend in to adopt SoC architectures for embedded vision. These SoCs are usually composed by an ARM microprocessor and at least an additional hardware component, which can be a FPGA, a GPU or a DSP. The low-processing part of the algorithm is run in the FPGA, GPU or DSP, and the rest in the microprocessor, combining the strengths of both architectures. Additionally, as they are physically located in the same chip, the overall power consumption is much lower than having them in two separate chips. Israeli company, Mobileye, the global leader in the design and development of camera-based Advanced Driver Assistance Systems has introduced its latest System-on-Chip (SoC) for use in multi-camera installations. Engineering samples of EyeQ4 are expected to be available by the fourth quarter of 2015, and series production is expected in early 2018. [20]. NVIDIA says its new Tegra K1 mobile processor will help self-driving autonomous cars advance from the realm of research into the mass market, with its automotive-grade version of the same GPU that powers the world’s 10 most energy-efficient supercomputers. The K1 features a quad-core CPU and a 192-core GPU using NVIDIA’s Kepler architecture, which forms the basis for the company’s range of powerful GPUs that are used in supercomputers. [21]

The new TDA2x SoC family, from Texas Instruments [22], with cutting-edge Vision AccelerationPac technology and XILINXC [23] with the portfolio SoC solutions for automotive industry are other players to be considered.

5.1.3 Embedded CPVS for Smart Building Management

Smart Facility and Building Management (SF&BM) generally involves a number of disciplines and services. The most general description to identify the market segment is understanding Smart F&BM as integrated management process that considers people, process and place in organisational context, being focused in the design and improvement of intelligent buildings (IB) and the coordination and optimization of several domains: facilities, life security, physical security and information technology. In this context, companies are becoming more interested in exploring opportunities to consolidate multiple services from single suppliers as a way of improving value. There is a significant consolidation opportunity for service providers able to deliver an integrated solution.

As described before, one of the key trends is to provide solutions that can take remedial actions automatically, providing a coordinated response in the “foundational systems” such as security, electrical and lighting distribution or HVAC (heating, ventilation, and air conditioning).

In Smart building domain, it is important to remark that heating, cooling and ventilation accounts for 30% energy usage and for 50% of the electricity. Currently, most modern buildings still condition rooms assuming maximum occupancy rather than actual usage. As a result, rooms are often over-conditioned needlessly.

In this sense, the video surveillance system, that is a mainstay of building security, may provide advanced information for energy management and overall building performance, like the proposed approach for HVAC control strategy based on occupancy prediction and real time monitoring.

However, though during last years a wide range of new applications within computer vision have been enabled, the network bandwidth, server processing and cost have been inhibitors for these opportunities up until now. Additionally, the traditional vision of a vertically structured market prevented the adaptation to a growing demand of dynamism and flexibility in the context of Smart Facility Management. The market is demanding not only more efficient, flexible and autonomous surveillance systems, but the integration of video systems to provide more data and information for energy management and enhanced building performance.

In this context, the Cognitive & Perceptive Video Systems (CPVS) enabled by COPCAMS would represent a significant step towards wider adoption of embedded vision systems within the smart facilities & smart building management domain. This new approach would provide advanced features in an emerging market, where and improved performance and reduced energy consumption will facilitate the use of embedded cameras not only as simple sensors, but as a distributed cognitive system, going beyond smart surveillance.

The target, at project end is to provide an integral approach (together with the use case in T5.1: Cognitive and Perceptive Cameras Systems for Smart Building Management Domain), in order to identify a minimum viable service (MVS) that can be provided to different clients as a comprehensive solution within the Smart Facilities and Smart Buildings Management domain, taking into account performance, cost, efficiency and deployment requirements.

5.1.4 Motion detection

The field of change detection within image processing and computer vision is extensive and well-established. However, according to a recent paper [27], a challenging dataset suitable for benchmarking has been lacking. The paper's authors describe a new dataset that they have compiled and that they deem sufficiently varied and substantial to allow comparison between change detection methods over the next few years. As well as introducing the new dataset, reference [27] also presents a survey of change detection techniques and proceeds to describe an evaluation of these techniques by means of the dataset.

The survey notes that change detection approaches can be categorised into: frame differencing; parametric and non-parametric background modelling; motion analysis; global image modelling; and machine learning. The evaluation assesses the extent to which regions of change are correctly segmented.

The performance evaluation documented in [27] ranks selected algorithms by integrating the scores for multiple metrics. The top ranking algorithms are diverse in terms of their core techniques. The top three methods as given by ascending rank employ respectively: models of surface reflectivity; Gaussian mixture models; and non-parametric probability distributions.

The paper notes that some individual metrics correlate better than others with the overall rankings. One such metric is the percentage of incorrect classifications given as:

$$\frac{(\text{false negative} + \text{false positive})}{(\text{all instances})} x 100\%$$

The three highest ranking methods return the following scores for this metric: 1.85%, 1.50% and 1.77%. True positive rates for these three methods are: 78%, 77% and 78%; the corresponding false positive rates are 0.8%, 0.6% and 1%.

5.1.5 Object detection and tracking

Algorithms for object detection in video are based on image pixels processing aimed at determining areas related to moving objects. The most often used approaches are based on background subtraction using algorithms such as Gaussian Mixture Models [28] or Codebook [29]. These algorithms are complex and the computation time depends on the video resolution. Since these algorithms are parallel in nature, they are suitable for implementation on multi-core and many-core platforms. It is possible to achieve a speed-up of these algorithms by means of parallel, multi-threaded implementation on a multi-core CPU platform, using e.g. OpenMP system. Experiments proved that CPU platforms are not powerful enough for processing high resolution video streams (below 20 fps for 1920x1080 video on Quad Core 2.5GHz CPU) [30]. Therefore, attempts were made to implement the background subtraction algorithms on many-core parallel platforms such as GPU [31] or FPGA [32]. As initial implementations proven successful and promising, the state of the art solutions were chosen for implementation and optimization in COPCAMS Project.

Object tracking algorithms use the detection results for tracking the movement of objects and to detect important events. Typical solutions, such as tracking based on Kalman filters are computationally simple, but they do not cope with object occlusion or fragmentation [33]. More advanced tracking algorithms are based on particle filters that model object parameters, e.g. color histograms [34]. Such algorithms increase the computation time significantly, but they provide much more accurate tracking results. Both the particle filter estimation and updating, as well as histogram calculation, may be implemented in parallel, using many-core computing platforms, which may allow for real-time video analysis. It should be noted that current implementations of algorithms require a powerful computer system. In practical situations, distributed systems composed of cameras, analyzing stations and a central server has to be built, allowing for distributed processing and handling the results [35]. A step beyond state-of-the-art is based on low-power device possible to be embedded in the camera, reducing the complexity of such a system.

Building smart cameras is the current prominent trend in video surveillance, although cameras available on the market are not equipped with parallel real-time processing units for object detection and tracking. Thus, the COPCAMS Project aims at providing proof of concept for low power, highly parallel platforms and optimized object detection and tracking algorithms.

5.2 Targets at project end

5.2.1 Acoustic Detection and Data Fusion Algorithms

The prototype should allow for accurate detection of the defined event and for determining and tracking the culprit. Thus following success criteria are defined, expressed as a percentage achieved for 60 repetitions (Table 2):

Factor	**Value**	**Comment**
correctly detected events, true positives	> 90%	both audio and RFID cues are detected
incorrectly detected events, false positives	< 2%	1 occurrence at most for all 3 hours of recordings
correctly determined culprit, true positives	> 95%	
correctly tracked person between cameras A and B, true positives	> 95%	for all 3 hours of recordings, including casual walk, idling, and empty scene
correctly tracked person between cameras B and C, true positives	> 95%	for all 3 hours of recordings, including casual walk and idling, and empty scene

Table 3 - Success criteria

5.2.2 Image Processing Algorithms

Three methods have been evaluated for the purpose of developing a modified RMD. Baseline results were used to benchmark these advances.

Increases in frame rate were attained by optimisation on the embedded platform. Once the initial port had been confirmed as equivalent to the reference implementation, the bottlenecks in the system were identified and optimisation applied to tackle these bottlenecks. The details of this process are covered in more detail in deliverable D3.5 [2].

6 Current state of the demonstrator

6.1 Building Surveillance Scenario

The prototype will be based on 3 hardware platforms, discussed previously in project deliverables, namely:

- Beagle Board will be used for RFID signals acquisition and pre-processing
- Beagle Board will be used for audio signals pre-processing (detection and angle of sound arrival measurement)
- Jetson TK1 will be used for video content analysis – object tracking and recognition, as well as event detection, using algorithms running on a parallel processor

GUT-MSD team developed required algorithms and methods for:

- multiple cameras communication, object tracking, and re-identification
- tracking improvements based on flow statistics modelled by Pawlak's Flow Graphs
- acoustic events detection, classification, and localization
- RFID tags localization and tracking by a multi-sector directional antenna

The reader is directed to following deliverables for details on hardware specifications, and algorithm documentation:

- D3.1 Image Analysis Methods & Algorithms – Platform Independent [3]
- D3.2 Video Algorithms – Platform Independent [4]
- D3.3 Innovative Multisensor Analysis Techniques – Interim Report [5]
- D4.1 COPCAMS Platform – Initial Hardware [8]
- D4.2 COPCAMS Middleware – Preliminary Version [9]

6.2 Traffic Surveillance Scenario

The deployment for traffic surveillance demonstrator has been done in a motorway near Tecnalia's facilities as is stated in 4.2. It is mounted in a gantry sign structure and it includes an IP camera with an environmental cover connected to a PC in the cabin near the road.

From this deployment, a set of videos have been recorded to test and compare the outputs between the different platforms.

The algorithms to be tested are reported in the documents [3, 4, 6, 7]. The selected platforms to be deployed these algorithms will be a standard PC and a GPU.

6.3 Building Surveillance Scenario (RMD)

6.3.1 PC-based

No Demonstrator was planned for TRT-UK, in the COPCAMS TA. The work conducted on PC-based experiments described in Section 3.3.1, formed the equivalent of a live demonstrator, since the datasets used were collected in environmentally challenging situations.

6.3.2 Embedded

No Demonstrator was planned for TRT-UK, in the COPCAMS TA.

Note: TRT-UK intend to demonstrate the ported RMD at the COPCAMS Gdansk review, November 2015.

6.4 Smart Facility and Smart Building Scenario

The preliminary studies and tests performed during lab experiments have allowed to identify the target platforms, candidate algorithms and define the specification for the field tests and validation strategy. The scope and final deployment of the use case is currently conditioned to the amendment requested by the Spanish consortium on July 2015.

7 Conclusion

In this document the different use cases and scenarios for the Indoor & Outdoor Surveillance Applications have been presented. They include a wide variety of cases covering current state of the art issues such as robust motion detection for any daytime, people tracking in buildings, incident detection in tunnels…

There is also a brief description of the different technologies that will be applied in the demonstrator at the project starting time (object tracking, traffic systems, motion detection…). It shows the state of the technology at that point and will be used to test how much improvement will be achieved during the project.

This document also includes how the different algorithms will be evaluated during the experiments. What is going to be evaluated in each platform and how. These algorithms have been described in previously documents being platform independent [3, 4, 5] and platform dependant [6, 7].

The results of these tests will be reported in deliverable D5.6 Outdoor & Indoor building Surveillance Applications Report.

References

1. ELAN. www.mpi.nl/corpus/manuals/manual-elan_ug.pdf
2. COPCAMS deliverable D3.5: Video Algorithms – On COPCAMS Platforms
3. D3.1 Image Analysis Methods & Algorithms – Platform Independent
4. D3.2 Video Algorithms – Platform Independent
5. D3.3 Innovative Multisensor Analysis Techniques – Interim Report
6. D3.4 Image Analysis Methods & Algorithms – On COPCAMS Platforms
7. D3.5 Video Algorithms – On COPCAMS Platforms
8. D4.1 COPCAMS Platform – Initial Hardware
9. D4.2 COPCAMS Middleware – Preliminary Version
10. J. Bier. Implementing Vision Capabilities in Embedded Systems. Berkeley Design Technology Inc., November 2012.
11. Andre R. Brodtkorb, Christopher Dyken, Trond R. Hagen, Jon M. Hjelmervik, and Olaf O. Storaasli. State-of-the-art in heterogeneous computing. Sci. Program., 18(1), January 2010
12. Electronics Publications. Video Surveillance Market: Global Forecast and Analysis 2011-2016. Available: http://www.electronics.ca.
13. Visiongain: Automotive Advanced Driver Assistance Systems (ADAS) Market 2014-2024
14. http://www.itsinternational.com/categories/detection-monitoring-machine-vision/features/transportation-applications-move-to-machine-visions-mainstream/
15. http://www.its.dot.gov/research/pdf/TechScan%20computervision%20ITSA.pdf
16. http://www.citilog.com/products/products.html
17. http://www.flir.co.uk/traffic/content/?id=66601
18. http://www.prnewswire.com/news-releases/global-automatic-number-plate-recognition-anpr-market-security-and-surveillance-vehicle-parking-traffic-management-toll-enforcement---trends-and-forecast-2014---2020-300015658.html
19. Aissaoui, R. , Menouar, H. ; Dhraief, A. ; Filali, F. ; Belghith, A. ; Abu-Dayya, A. Advanced real-time traffic monitoring system based on V2X communications, Communications (ICC), 2014 IEEE International Conference on.
20. http://www.traffictechnologytoday.com/news.php?NewsID=66921
21. http://www.traffictechnologytoday.com/news.php?NewsID=55129
22. http://www.ti.com/product/TDA2
23. http://www.xilinx.com/publications/prod_mktg/ZynqAuto_ProdBrf.pdf
24. Al Akkoumi M., Huck R., Sluss J; A Personnel Detection Algorithm for an Intermodal Maritime Application of ITS Technology for Security at Port Facilities
25. Bilgic B., Horn B., Masaki I. Fast Human Detection with Cascaded Ensembles on the GPU

26. L. Chen, F. Chen, and X. Guan. A Video-Based Indoor Occupant Detection and Localization Algorithm for Smart Buildings
27. N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, ‘A Novel Video Dataset for Change Detection Benchmarking’, Image Processing, IEEE Transactions on, vol. 23, no. 11, pp. 4663–4679, 2014.
28. Stauffer, C., Grimson, W.E.: Adaptive background mixture models for real-time tracking. Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 246–252, USA (1999)
29. Kim, K., Chalidabhongse, T.H., Harwood, D., Davis, L.: Real-time foreground-background segmentation using codebook model. Real-time Imaging. 11, 167–256 (2005)
30. Szwoch G. Ellwart D., Czyzewski A.: Parallel implementation of background subtraction algorithms for real-time video processing on a supercomputer platform, Journal of Real-Time Image Processing, DOI 10.1007/s11554-012-0310-5 (2012)
31. Pham V., Vo P., Hung V.T., Bac L.H., GPU implementation of Extended Gaussian mixture model for Background subtraction, 2010 IEEE RIVF International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF) (2010)
32. Genovese M., Napoli E., De Caro D., Petra N., Strollo A.G.M., FPGA Implementation of Gaussian Mixture Model Algorithm for 47 fps Segmentation of 1080p Video, Journal of Electrical and Computer Engineering, DOI: 10.1155/2013/129589 (2013)
33. Szwoch G., Dalka P., Czyzewski A., Resolving conflicts in object tracking for automatic detection of events in video, Elektronika 52(1), pp. 52-55 (2011)
34. Isard M., Blake A., CONDENSATION—Conditional Density Propagation for Visual Tracking, International Journal of Computer Vision, 29(1), pp 5-28 (1998)
35. Czyżewski A., Szwoch G., Dalka P., et al, Multi-stage video analysis framework, in: Video surveillance, W. Lin, Ed,. InTech, pp. 147-172 (2011)