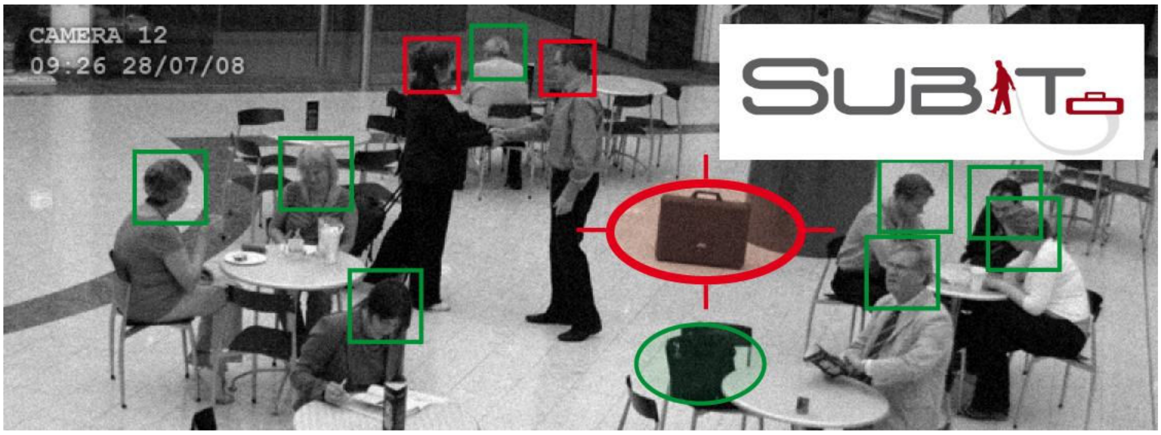
FP7

Seventh Research Framework Programme



Grant Agreement No. 218004
SUBITO

Surveillance of Unattended Baggage and the Identification and Tracking of the Owner

SUBITO Deliverable D100.2: Final Report

Due date: December 2011

Project co-ordinator: Mike McPartlin
Project Manager
SELEX Galileo

Tel: +44 (0) 131 343 5204
Fax: +44 (0) 131 343 8110
E-mail: mike.mcpartlin@selexgalileo.com
Web: www.subito-project.eu

Issue 2

CLASSIFICATION:	**PP (Restricted)**

EXECUTIVE SUMMARY

The Surveillance of Unattended Baggage and the Identification and Tracking of the Owner, (SUBITO), project was a research & development project funded by the European Commission Framework 7 programme. Classified as a European Union capability project it was designed to research and further develop novel technology to achieve a set of top level objectives, namely:

- Fast detection of baggage that has been abandoned,
- Fast identification of the individual who left the baggage, and
- Fast determination of their location or path they followed.

The project was guided by end-user driven requirements to ensure that security personnel receive the technologies they need in order to deliver improved threat security.

The system requirements were derived from an understanding of the current user perspective to unattended baggage threat scenarios and the logic applied to identify alarm and non-alarm conditions. This allowed a system architecture to be developed which was scalable to any size of installation and would be able to perform with the key measures which a fielded system should attain in a real application. A theoretical Privacy Impact assessment was performed on the developed architecture and recommendations made as to how privacy issues be addressed in future work in this area. The wider social and legal aspects of the technology have also been studied.

The development of the system architecture was supported by a series of additional studies which analysed the benefits to system function and performance to be gained by the use of improved camera technology, additional sensors or distributed processing schemes. The outputs from these studies influenced the generic system architecture design to ensure that these technologies could be implemented in to a future systems while maintaining a scalable architecture.

The implementation of the system architecture required both image analysis and threat assessment algorithms to be developed. The key goal of the image analysis algorithms was to develop the capability to robustly detect, segment, track and classify moving objects within the scene. This was achieved by using a multi-view approach which has reduced system false alarms, and produced clean detections. In addition, both track closeness and recall are improved, a quality which benefitted the performance of the threat analysis algorithm. Further tests to improve the image analysis algorithm performance and robustness are on going with specific interest in transferring the current software implementation to graphical processing units to offer a considerable speed increase.

The threat assessment algorithms were developed to classify potentially critical situations, given positional and classification data about the objects and people within the sensed environment. The experimental results achieved demonstrated that the inclusion of reasoning about the intentions of individuals within a scene and the interactions between these individuals leads to greatly improved performance over the state of the art in the detection of the threat of abandoned baggage. In particular, the performance of the developed SUBITO system exceeds that achieved in the previous ISCAPS study.

The project culminated in a final demonstration and evaluation of an integrated system operating against pre recorded scenarios, as defined during the system requirements phase, designed to show capability against the top level objectives mentioned above. A parallel workshop allowed discussion of the project results and future prospects in several areas of the implemented technology solutions.

DOCUMENT CHANGE

Date	Author	Issue	Subject of Change
21st Dec 2011	P. C. Findlay	1	Initial Issue
26th Sept. 2012	P. C. Findlay	2	Update to Exploitation Plans

DISCLAIMER

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 218004

The views expressed in this document are those of the author(s) and the European Community is not liable for any use that may be made of the information contained herein.

CLASSIFICATION:	**PP**

TABLE OF CONTENTS

1 INTRODUCTION 1

1.1 Document Purpose and Scope 1

1.2 The SUBITO Project 1

1.3 SUBITO Consortium 1

1.4 Project Background and Objectives 2

1.5 Project Structure 4

2 DESCRIPTION OF MAIN SCIENTIFIC RESULTS/FOREGROUNDS.........................5

2.1 Scenario Definition 5

2.2 System Definition 6

2.3 System Architecture 7

2.4 System Test and Demonstration Planning 10

2.5 System Algorithm Development 11

3 INTEGRATION, TESTING & DEMONSTRATION 26

3.1 Integration 26

3.2 Testing 28

3.3 Results 29

3.4 Final Demo 31

3.5 Competitive Assessment 32

3.6 Exploitation plans 34

4 DESCRIPTION OF POTENTIAL IMPACT AND DISSEMINATION 43

5 PUBLIC WEBSITE AND CONTACT DETAILS 44

LIST OF FIGURES

Figure 1-1: SUBITO Project Work Breakdown Structure 4
Figure 2-1: System Context Diagram 6
Figure 2-2: Other Sensors Study Functional Categories 8
Figure 2-3: SUBITO Sensor / Functional hierarchy 8
Figure 2-4: SUBITO System Base Configuration 10
Figure 2-5: a) Image and b) Threat Analysis Logical Architectures 11
Figure 2-6: Example of person detection in multi-view scene 12
Figure 2-7: Example of bag detection in multi-view scene 12
Figure 2-8: PTZ Camera Sub-system. 13
Figure 2-9: Artefact removal using Spatial Reasoning 15
Figure 2-10: Facial Detection using single probe image 17
Figure 2-11: Facial Detection using multiple probe images 17
Figure 2-12: Image Patch Extraction from PTZ Cameras 18
Figure 2-13: Multi-view Object Matching. 19
Figure 2-14: Snapshots from Threat Assessment Test Sequences 20
Figure 2-15: Colour-coded trajectories between waypoints, leading to destination 21
Figure 2-16: Illustration of Predicted versus Tracked Ground Truth Trajectories 22
Figure 2-17: Main SUBITO MMI Window & Full resolution camera display 25
Figure 3-1: Implementation Architecture 26

LIST OF TABLES

Table 1-1: SUBITO Consortium Members 2
Table 2-1: Results for *Ownership* + *Social Groups* Definition. 24
Table 3-1: Summary of the performance on 15 different sequences 30
Table 3-2: End User feedback on SUBITO principle capabilities 31
Table 3-3: How SUBITO meets End User requirements. 32
Table 3-4: Top level functionality comparison 33
Table 3-5 : Future FP7 Programme Involvement 34

ACRONYMS

3D	3-Dimensional
CCTV	Closed Circuit Television
CPU	Central Processing Unit
EC	European Commission
EU	European Union
FoV	Field of View
GPU	Graphics Processing Unit
GT	Ground Truth
GUI	Graphical User Interface
i-LIDS	Image library for intelligent detection systems
IED	Inter Eye Distance
ISCAPS	Integrated Surveillance of Crowded Areas for Public Security
MHT	Multiple Hypothesis Tracker
MMI	Man-Machine Interface
PC	Personal Computer
PETS	Performance Evaluation of Tracking and Surveillance
PIA	Privacy Impact Assessment
PTZ	Pan-Tilt-Zoom
R&T	Research & Technology
SAD	System Architecture Document
SFM	Social Force Model
SLE	Social, Legal and Ethical
SQL	Structured Query Language
SRD	System Requirement Document
SUBITO	Surveillance of Unattended Baggage and the Identification and Tracking of the Owner
UB	Unattended Baggage
WLAN	Wireless Local Area Network
WP	Work Package
XML	eXtensible Markup Language

DEFINITIONS

Alarm	*The automatic reporting by the System of an event to the operator.*
Baggage	*Any object that can be carried by hand and that can be used to contain other objects.*
Consortium	*The group of large companies, Small & medium Enterprises, Universities and Research & Technology Organisations comprising the project team.*
Customer	*European Union.*
End User	*Security personnel, system operator and maintainer (if any) who uses the system.*
Event	*Detection of Unattended Baggage.*
Identification	*The act of associating a particular individual with an item of baggage whose status has triggered the unattended baggage event.*
Operator	*An end user person or persons responsible for operating and monitoring the CCTV system.*
Owner	*The individual who enters the scene of interest with the baggage.*
Real Time	*The rate of system operation (defined during the project research).*
Scene	*Surveyed area with zero or more people.*
Stakeholder	*A representative of an Organisation interested in the systems overall security capability and technology.*
Surveillance	*Defined to be a collection of video imagery from a CCTV network (consisting of 1 or more cameras) arranged in zones.*
Suspicious Event	*An action or object displaying questionable behaviour.*
Suspicious Person	*An individual displaying questionable behaviour or engaged in questionable activities.*
System	*The SUBITO demonstration system.*
Tracking	*Observations used to form and maintain a continuous estimate of the location of a designated individual, or object, with respect to one or more reference frames.*
Unattended	*Object defined to be unattended when separated from its owner by a specific distance. Further, the item of baggage is defined to have been abandoned when left unattended by the owner for a defined period of time, during which the owner (or second party) has not re-attended to the baggage.*

REFERENCES

1. SUBITO Proposal Annex I, Version 1.0, 19th August 2008.
2. SUBITO Final Report Annex, *Annex to Project Deliverable D100.2*, **Issue 1**, July 2011.
3. SUBITO System Requirements Document, *Project Deliverable D200.1*, **Issue 1**, April 2009.
4. SUBITO Baseline Test & Demonstration Plan, *Project Deliverable D200.2*, **Issue 2**, March 2011.
5. SUBITO System Architecture Document, *Project Deliverable D200.3*, **Issue 1**, April 2010.
6. SUBITO Privacy Impact Assessment, *Project Deliverable D200.4*, **Issue 1**, June 2011.
7. Use of Other Cameras, *Project Deliverable D300.1*, **Revision 2**, February 2010.
8. Use of Other Sensors, *Project Deliverable D300.2*, **Issue 1**, February 2010.
9. Distributed Targeting Report, *Project Deliverable D300.3*, **Issue 1**, February 2010.
10. Ethical and legal aspects of SUBITO, *Project Deliverable D300.4*, **Issue 1**, May 2011.
11. Threat Analysis Algorithms Technical Report, *Project Deliverable D400.1*, **Issue 1**, February 2011.
12. The Ethics of Automated Threat Assessment, *Project deliverable D400.2*, **Issue 1**, February 2011.
13. Image Analysis Algorithms Technical Report, *Project Deliverable D500.1*, **Issue 1**, February 2011.
14. SUBITO Interface and Data Management Specification for the Demo System, *Project Deliverable D600.1*, **Issue 2**, March 2011.
15. Description of Integrated System, *Project Deliverable D600.2*, **Issue 1**, February 2011.
16. Database Archive of Collected Video, *Project Deliverable D600.3,* **Issue1a**, February 2011.
17. SUBITO Initial System Operation and Pre-demonstration, *Project Deliverable D600.4*, **Issue 1**, February 2011.
18. SUBITO Grant Agreement, *No. 218004*, 15th December 2008.
19. SUBITO Consortium Agreement, **Issue 2**, 11 November 2010.
20. SUBITO Progress Report 1, *Project Deliverable D100.1b*, **Issue 2**, January 2010
21. SUBITO Progress Report 2, *Project Deliverable D100.1d*, **Issue 2**, January 2011
22. SUBITO Progress Report 3, *Project Deliverable D100.1f*, **Issue 1**, October 2011
23. C. Guillot, et al. “Background subtraction for PTZ performing a guard tour and application to cameras with very low frame rate” *Tenth International Workshop on Visual Surveillance in conjunction with ACCV*,2011.
24. W. Ge, et al. “Automatically detecting the small group structure of a crowd”, *Applications of Computer Vision 2009 Workshop on*, 1–8, (December 2009). *doi: 10.1109/WACV.2009.5403123*
25. S. Pellegrini, et al. “Improving data association by joint modelling of pedestrian trajectories and groupings”, *Computer Vision ECCV 2010*, **6311**, 452–465. Springer, (September 2010). *ISBN: 978-3-642-15548-2*
26. Interviews Reports on Situational Awareness, *Project Memo MEM_EXT_0377*, **Issue 1**, December 2011.

27. SUBITO System Test & Evaluation: WP650 Requirements and Rational, *Project Memo MEM_EXT_0339*, **Issue 1**, March 2011
28. SUBITO – Evaluation Results, *Project Report RPT_EXT_0361*, October 2011
29. Report on Final Demo & Review, *Project Deliverable D600.5*, **Issue 1**, November 2011.
30. L.Usai, F.Pantisano, L.Roncarolo, L.Vaccaro, F.Selvaggi, "SuperVision: Video Content Analysis Engine for videosurveillance applications", chapter .. INTECH …. , 2010
31. SUBITO Final Report, Online Version, *Ref: intermediateReport934106*, **Issue 1**, Dec 2011.

1 INTRODUCTION

This document is the Final Report for the Surveillance of Unattended Baggage and the Identification and Tracking of the Owner (SUBITO) project. The final report forms project deliverable D100.2: Final report to the European Commission (EC) as detailed in Ref. 1.

1.1 Document Purpose and Scope

This report provides a summary of the work which has been carried out by the SUBITO consortium over the 34 month duration of the project. It includes a discussion on the wider social impact of the work and reports the dissemination activities that have taken place to inform the wider audience about the SUBITO project. A more detailed overview of the work can be found in the supporting Annex to this document, [Ref. 2], where the individual consortium members have contributed to the relevant work packages.

Details of the sub-system development testing and evaluation can be found in the associated documentation, [Refs. 3 - 17].

1.2 The SUBITO Project

SUBITO was a Research and Development (R&D) project funded by the EC Framework 7 programme under grant agreement no 218004, [Ref. 18]. It was classified as an European Union (EU) capability project and was designed to research and further develop novel technology for automated, real time detection of unattended baggage and fast identification of the individual responsible (the 'Owner'), their subsequent path and current location.

The SUBITO project was required to deliver at an end user site a demonstration of semi-automated, man-in-the-loop data processing operating with existing Closed Circuit Television (CCTV) technology and demonstrating:

- Fast detection of baggage that has been abandoned,
- Fast identification of the individual who left the baggage, and
- Fast determination of their location or the path they followed.

1.3 SUBITO Consortium

The SUBITO consortium was formed from a diverse group of technology and implementation experts from across the EU and involved leading representatives from Industry, the Research and Technology (R&T) community, Academia and End Users, see Table 1-1.

The large number of partners from six countries allowed the project to gain benefit from the national experiences across the European Union.

The details of the agreements forming the basis of the cooperation of Companies and Institutions inside the SUBITO Project can be found in the Consortium Agreement [Ref.19].

Table 1-1: SUBITO Consortium Members

Consortium Member	Member Type	Country
SELEX Galileo Ltd	Industry	UK
ELSAG DATAMAT SpA	Industry	Italy
Office National d'Etudes et de Recherches Aérospatiales	R&T	France
L-1 Identity Solutions AG	Industry	Germany
Commissariat à l'énergie atomique	R&T	France
University of Leeds	Academic	UK
University of Reading	Academic	UK
VTT Technical Research Centre of Finland	R&T	Finland
Austrian Institute of Technology	R&T	Austria
Fiera di Genova SpA	End User	Italy
University of Oxford[1]	Academic	UK

1.4 Project Background and Objectives

The problem of 'Unattended baggage' detection was part of the scope of a previous EU project called Integrated Surveillance of Crowded Areas for Public Security (ISCAPS). As part of this project, a technology road map had been produced which captured the vision for future research and development required for the technology to meet the operational requirements of potential End Users for such a system. The main outcome from the roadmap was a set of basic requirements for a surveillance system such as SUBITO:

- Wide area coverage by the cameras,
- Potential inclusion of smart sensors,
- Day, night, all weather operation,
- Person Identification, and
- Minimisation of false alarms

In addition ISCAPS demonstrated that image processing capabilities exist that can identify and track goods and individuals within a simple scenario, and could be used to detect isolated goods which have been abandoned for a defined period of time. However, the existing approaches had difficulty in complex and cluttered scenes, resulting in high false alarms or missed detections, especially in the case where the owner leaves goods within the vicinity of other people without their knowledge or cooperation. Detection of such threatening situations requires the explicit introduction of the concept of "ownership" and identification of the owner.

[1] University of Oxford joined the consortium in December 2010 to replace the Legal and Ethics Expert lost from University of Reading.

The SUBITO project aimed to exceed ISCAPS 'unattended baggage detection' capability by including abandoned baggage owner identification and tracking and produce a demonstration system that would showcase this capability.

Starting from current state-of-the-art technology, the SUBITO project was tasked with developing solutions beyond those that currently existed, that would meet a number of key scientific and technical objectives designed to demonstrate the capability of the system in the task of detecting unattended baggage.

These basic requirements listed above from ISCAPS and the need to identify and track the owner formed the basis for the key scientific and technological objectives for SUBITO, as stated in Ref. 1:

Objective 1. Find abandoned luggage and identify and track the owner.

Objective 2. Reduce the number and impact of false alarms.

Objective 3. Demonstration of automated detection of abandoned goods, fast identification of owner who left them and fast determination of the owner's location or their path.

Objective 4. Demonstrate a scalable route to implementation.

Objective 5. Examine the wider user of technologies for explosive threat identification in this context.

Objective 6. Examine the use of Pan-Tilt-Zoom (PTZ) camera technologies to distinguish between threatening and non-threatening goods.

Objective 7. Manage/improve public perception of this technology and its implications.

To meet these objectives, the SUBITO project aimed to advance technology beyond state-of-the-art in areas such as:

- Abandoned Object Detection and Tracking of Owner,
- Robust Detection,
- Robust and Long Term Tracking,
- Robust Identification,
- Behavioural Analysis,
- Facial Recognition, and
- PTZ cameras.

In addition the project aimed to consider the wider aspects related to semi-automated surveillance systems by looking at the specific privacy issues associated with automated threat detection as well as the wider legal and ethical impacts of the technology particularly where elements are technically possible but which would not be socially or legally acceptable.

An important part of the SUBITO project was to ensure that the development of the system was targeted to meet End User needs. To this end the project consulted a variety of End Users during the requirement and scenario definition stage and sought End User feedback from the final demonstration of the developed system.

1.5 Project Structure

The SUBITO project was structured as shown in Figure 1-1.

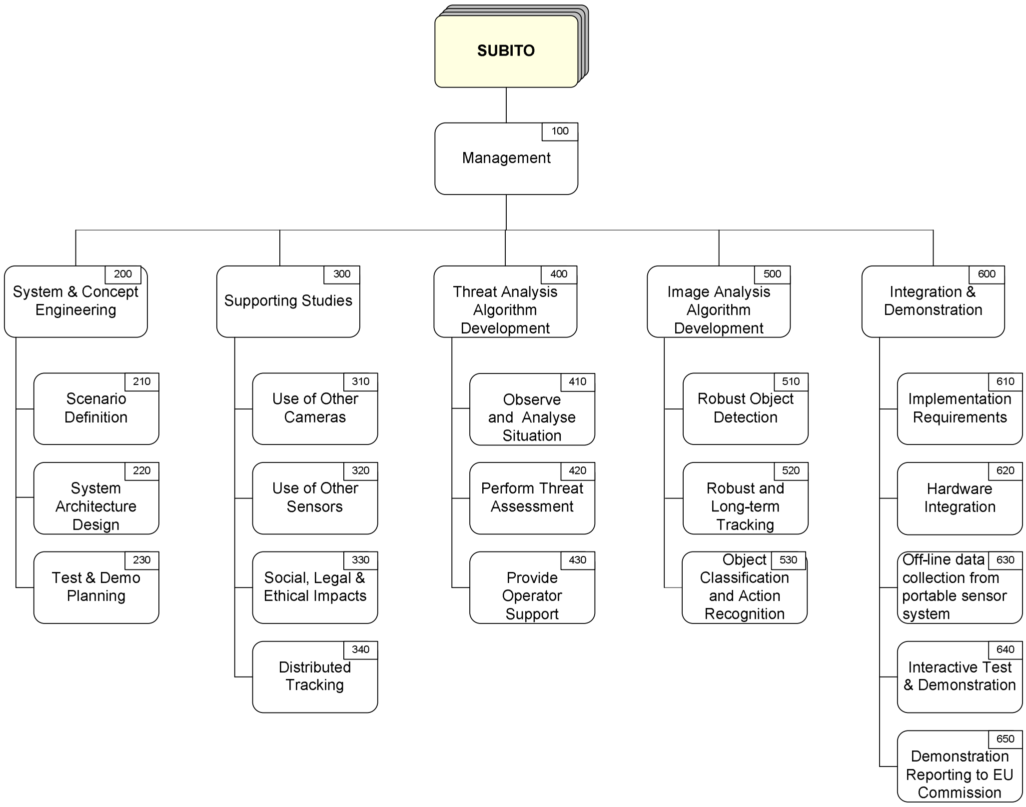


Figure 1-1: SUBITO Project Work Breakdown Structure

Overall management and co-ordination of the consortium partners was undertaken under Work Package (WP) 100.

WP200 focussed on generating the key requirements of the SUBITO project in consultation with End Users who had experience of public and private security, and in particular the issue of unattended and abandoned baggage. Once defined the requirements were used to design a baseline system architecture which would allow the development of a demonstration system to highlight the main capabilities of the SUBITO system within a representative environment, while also providing an architecture that was scalable to that required by any future commercial system.

The defined architecture, [Ref. 5], provided an initial point from which the algorithm development could take place. The main functions of the system algorithms were split into two areas, Image Analysis (WP500) and Threat Assessment (WP400). These were further split into sub-system elements which covered object detection, tracking and classification as well as observation, analysis and threat assessment.

As well as the main algorithms a Man Machine Interface (MMI) was developed as part of WP400. This allowed the main SUBITO system operating on representative scenarios to be demonstrated to End Users. End Users were asked to provide feedback on the system capabilities and functions.

In addition a number of supporting studies were carried out in WP300 to investigate the potential benefits to be gained through the use of enhanced CCTV, other sensors (such as acoustic, chemical etc.) and distributed tracking in the system. The issue of privacy and the legal and ethical aspects of the system were also considered.

2 DESCRIPTION OF MAIN SCIENTIFIC RESULTS/FOREGROUNDS

The following sections contain details on the main scientific progress made during the SUBITO project. The discussion will broadly follow the make-up of the project Work Breakdown Structure, Figure 1-1, with the exception of WP300 where the results and outcomes of the supporting studies will be discussed in relation to the tasks of the main project that they support rather than in isolation.

A discussion of the main project management tasks covered in WP100 can be found in Refs. 20, 21 & 22.

2.1 Scenario Definition

The initial task for the SUBITO project was to understand the threat and non-threat situations relating to unattended baggage from the perspective of the End User and how these situations would lead to the corresponding alarm and non alarm conditions. To achieve this, a number of End User visits were carried out to gain an appreciation of the problems faced by the End Users and to gain an understanding of a typical CCTV surveillance installation.

In parallel the knowledge gained during the previous ISCAPS project was drawn on to produce a view on the threat scenarios within the consortium. Both the ISCAPS background and End User consultations provided a view on the types of scenarios required to be covered by the project to ensure the demonstration and test encompassed genuine scenarios.

As a rule the SUBITO system was expected to perform under two general categories of scenario, S0 and S1, [Ref. 3]:

S0 – No alarm to operator.

In this scenario the baggage owner enters the scene, leaves baggage within the footprint of one or more ‘system’ cameras, stays within baggage location then leaves scene with the baggage.

S1 – Alarm to operator.

In this scenario the baggage owner enters the scene, leaves baggage within the footprint of one or more ‘system’ cameras, leaves baggage unattended or attended by a 3rd party and does not return within a defined period.

The complexity of these categories depended on the number of additional individuals and baggage objects that appeared in the scene. This information fed in to the baseline test and demonstration planning and informed the types of SUBITO specific scenarios captured for testing and validating system performance.

2.2 System Definition

The End User consultations provided input into the system design, resulting in the top level context diagram for the demonstration system as shown in Figure 2-1. The system was required to interface to an existing CCTV network comprising a number of imaging sensors [cameras] and a single system operator. For the demonstration system, the number of sensor interfaces was limited to a maximum of four, but this was not a fundamental limit as the system context was designed to be scalable to accept any number of sensor inputs. The purpose of the operator interface (MMI) was to allow an operator to view the raw and processed data from all cameras and control the operation of the data processing.

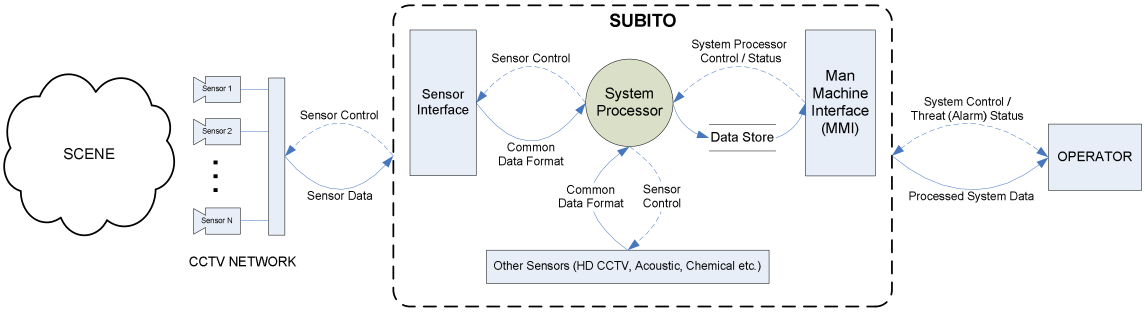


Figure 2-1: System Context Diagram

Figure 2-1 highlights the Sensor Interface, Man-Machine Interface (MMI), Data Store, Other Sensors (if available) and System Processor components of the system. These components provide the following system capabilities.

- The Sensor Interface would provide data acquisition support for each of the 4 camera interfaces to allow simultaneous data capture for processing by the system processor.
- The System Processor would operate on the camera data and provide the 'unattended baggage detection' and fast 'owner identification' and 'tracking' functions of the system.
- The MMI would interface the operator with the system processor and provide raw, processed and stored data visualisation and control facilities to enable the operator to control and demonstrate the systems capabilities.
- The Data Store would provide 'bulk, long term memory' storage for the system to support recording and playback of raw and processed video data to support 'unattended baggage' event analysis technology demonstration.
- If available the Other Sensors component would integrate any beneficial sensor technologies identified during the supporting studies work with the system processor for demonstration.

In addition the process of determining the system requirements, allowed certain Key Performance measures to be defined. Specifically it is *necessary* that for an effective fielded system to be adopted by end users, the performance should provide:

- Almost 100% 'suspicious' Unattended Baggage (UB) and owner event detection, low false alarm rate.

To maximise system effectiveness there is also a set of *desirable* Key Performance measures, presenting a more challenging design task:

- 100% 'suspicious' UB and owner event detection, with 0% false alarms.

A System Requirement Document (SRD), [Ref. 3], was created which captured key system requirements and performance criteria as detailed above.

2.3 System Architecture

With the key system requirements defined, the main focus switched to the system architecture, the design of which would allow a demonstration system that would show the capabilities of the SUBITO system in effectively detecting unattended baggage events, tracking the owner and presenting these to an End User.

A fielded SUBITO system would most likely have its front and back end fixed – i.e. the existing CCTV cameras in operation and the existing operator controls. As such the architecture design, and hence algorithmic design, was defined such the system would operate using the available data albeit with a stated performance limitation due to reduced resolution / lower frame rate etc. than a complete front to back SUBITO installation.

The primary drivers for CCTV systems can be distilled down to be basic requirements for camera parameters such as resolution, focal length etc. which are necessary to achieve acceptable performance. System modelling performed during the architecture design showed that while the basic tasks of baggage detection and individual tracking were feasible, it was not possible to achieve the required level of identification performance while maintaining as large an area as possible of basic surveillance coverage.

As a result of this modelling it was determined that the systems architecture for SUBITO should include one or more additional cameras operating in a narrow field of view (FoV) mode, e.g. a PTZ camera, whose purpose was dedicated to collecting high resolution imagery for use with facial recognition or other identification level algorithms.

To provide a comprehensive architecture design supplementary studies were performed to consider how other sensors and distributed processing could be incorporated into the system architecture to enhance the SUBITO system capabilities. A brief discussion of the outcome of these studies is provided below.

2.3.1 Use of other sensors study

The study into the use of other sensors defined a top level set of functionalities as shown in Figure 2-2, the purpose of these functional categories was to provide a general set of criteria against which different sensors could be assessed for down selection for more in depth analysis. This was required due to the large number of sensors types that exist.

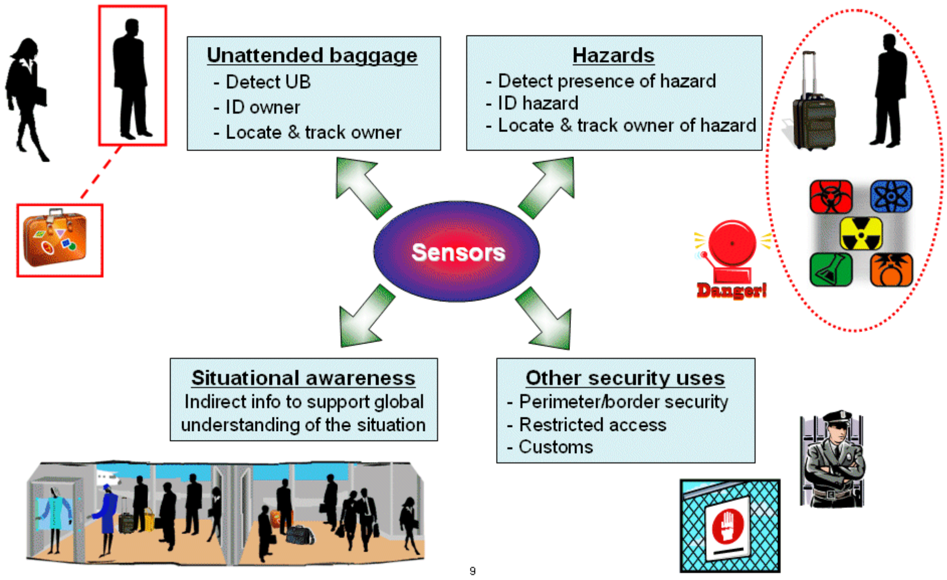


Figure 2-2: Other Sensors Study Functional Categories

The results of the top level systems design trade study showed that none of the individual sensors considered were currently capable of providing all the functionality required by SUBITO, [Ref. 8]. Furthermore some other sensor technologies are currently too immature to provide benefit within the SUBITO timescales. With the exception of a Hyperspectral system any sensor used would have to be used in conjunction with existing CCTV in order to be effective in detecting unattended baggage.

The use of a particular sensor within a SUBITO type system depends on the nature of the information they can provide. Some sensors, such as explosive threat identification systems would provide supporting information but would not provide a significant improvement on the primary functions of SUBITO obtained using CCTV.

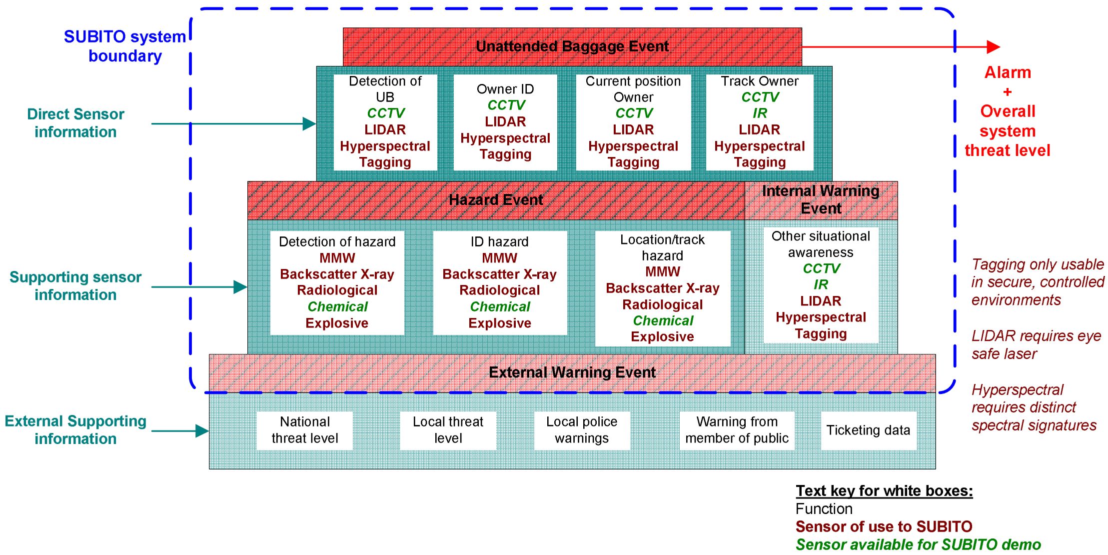


Figure 2-3: SUBITO Sensor / Functional hierarchy

Figure 2-3 summarises the findings of the study. Here the SUBITO functional hierarchy leading to the system providing an alarm event to the operator is shown. The types of other sensors investigated are mapped to this hierarchy, identifying the types of information and event that each can supply.

The SUBITO architecture takes into account the output of the additional sensor studies and makes provision for the inclusion of other sensors in future development, however for the final demonstration it was not considered likely that the identified sensors could be sourced and integrated alongside the core technologies.

2.3.2 Distributed system architecture study

In parallel to the other sensors investigation a study into distributed system architecture was undertaken which considered the practical aspects of working with a distributed system to perform the SUBITO task. The general aims were to identify the appropriate structure, carry out a cost-benefit analysis and define the basic requirements for this type of distributed system.

In a centralized system, such as the system context described above, section 2.2, where high resolution cameras are used, and data is generally received in a compressed format. This data must be decompressed prior to analysis which places a heavy loading operation on the processing, but the cost of using specialised hardware to ease the loading on the central processor is typically not very critical.

In comparison for a distributed system, low cost, high performance processors and dedicated decompression modules are required leading to considerably increased costs.

From the wireless transmission tests performed in the study, it is clear that for a distributed network of surveillance processors located in crowded areas where there are potentially many other Wireless Local Area Network (WLAN) users. The possibility exists for the communication links to deteriorate (either through heavy load of users or deliberate jamming). This is intolerable for potentially safety critical system.

The use of high gain directional antennas within the system design, could improve this situation for fixed long distance links; however, for such links a high level of external RF noise is possible which may interfere with system performance. Further details of the study outcomes as well as proposals for a proper distributed system architecture can be found in Ref. 9.

The study concluded that both compressed and decompressed image formats are need within surveillance systems. The compression/decompression process using software to achieve the modern compression standards overloads typical consumer Personal Computer (PC) processors.

In terms of the architecture for the SUBITO project and demonstration, no wireless or distributed processing was included due to the potential impacts on system performance and increased work in maintaining system security.

2.3.3 SUBITO Demonstration Architecture.

Based on the architecture design and supporting studies, the configuration of the demonstration was decided to be composed of 3 Fixed FoV CCTV sensors and 1 additional PTZ operating in a narrow FoV mode to provide data for identification purposes, Figure 2-4.

In this concept, entrance & exit points of the surveillance zone could also be equipped with additional cameras to provide high resolution imagery of the faces of individuals entering the scene.

Full details of the defined system architecture can be found in the Systems Architecture Document (SAD), [Ref. 5].

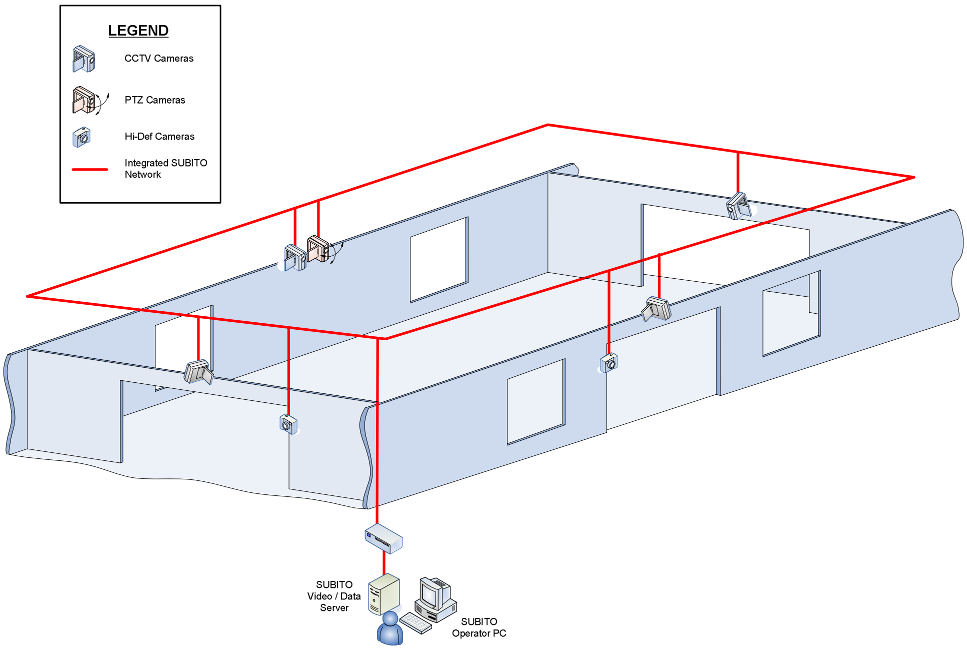


Figure 2-4: SUBITO System Base Configuration

2.4 System Test and Demonstration Planning

The definition of the system requirements and initial drafts of the system architecture allowed a baseline test and demonstration plan to be developed. The aim was to provide a working implementation of the SUBITO system that could be tested under as many representative scenarios as possible.

Based on a review of the system requirements, the testing would be carried out using commercial equipment wherever possible, including cameras, computational equipment and networks. The actual testing approach comprised defining a series of scenarios, carefully constructed to allow repeatability and easy of measure of performance.

The basic scenarios defined above, see Section 2.1, were expanded to create a series of test sequences which increased in complexity and which reflected the reality of baggage abandonment and would allow verification of the SUBITO stated objectives. The complexities of the individual sequences are graded difficulty: from single person clean area abandonment to those taking place in the middle of a group of people carrying baggage. In each case the sequence content was checked to ensure that all privacy requirements and safety constraints had been met prior to being recorded at the test and demo site.

Full details of the baseline test and demonstration planning can be found in Ref. 4 while a complete record of the collected video can be found in Ref. 16.

2.5 System Algorithm Development

The main algorithm work of the SUBITO project was split into two areas, image analysis and threat assessment, see Figure 2-5.

Within the SUBITO contexts, three types of image analysis algorithms were addressed: object detection, object tracking and object classification. The threat assessment algorithms were split into the areas of observe & analyse and threat classification, these will be discussed later.

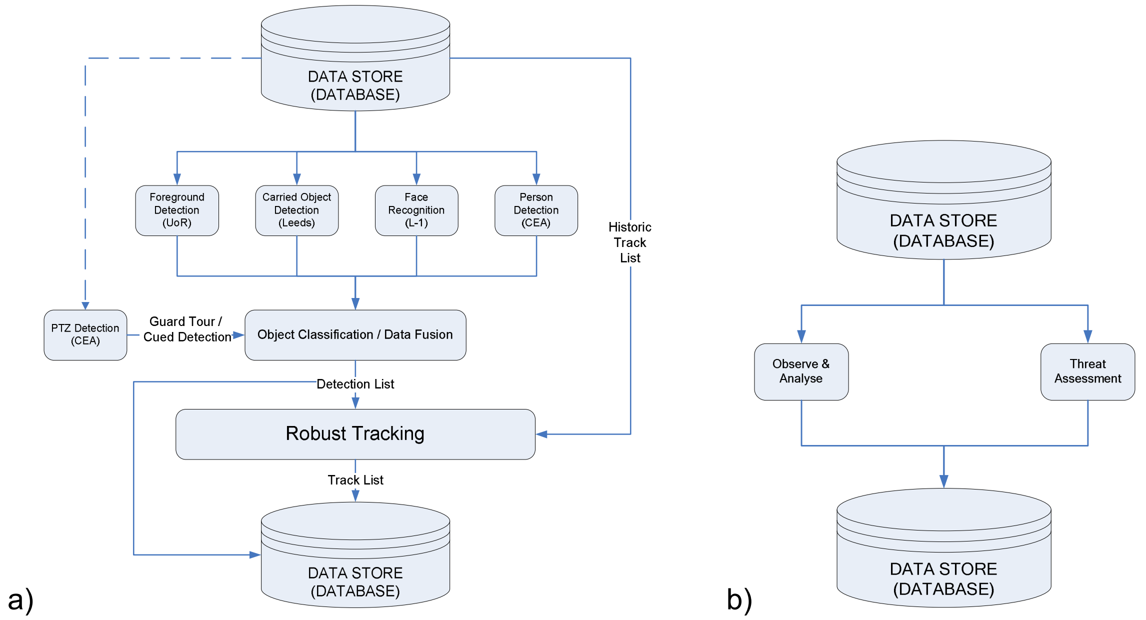


Figure 2-5: a) Image and b) Threat Analysis Logical Architectures

The following sections provide discussion on the work carried out in each of the algorithm development areas.

2.5.1 Object Detection

The first algorithmic step within the Image Analysis process was object detection. The object detection algorithms were developed incrementally over the course of the SUBITO project, with the performance of the algorithms tested at each iteration using representative data collected from the test site, and the results fed back in to the development process.

Initially a background and static object detection algorithm was implemented, which detected static regions within the scene, i.e. baggage hypotheses, which were validated against various criteria to report potential abandoned baggage. The next phase improved the robustness of the baggage verification and developed a multi-view detection algorithm that would allow the detected object to be positioned in 3D space. The final phase implemented an enhanced baggage verification algorithm and the shift of the algorithm from a purely Central Processing Unit (CPU) based algorithm to an alternative Graphical Processing Unit (GPU) based algorithm which provided a considerable speed improvement.

A number of component state-of-the-art algorithms were used in the object detection to provide robust object segmentation, specifically the detection of people and baggage within the scene. These were designed to provide robust background subtraction, static object detection and probable 3D location in space, utilising adaptive processes to automatically adapt to the current scene, and provide resilience against changes in the image, e.g. lighting levels. A fusion of detection from multiple views was also utilised to create a single world view of the likely location of detected objects.

This multi-view approach reduced false alarms, [cf. Objective 2], and produced robust detections of individuals, Figure 2-6, and of baggage, Figure 2-7.

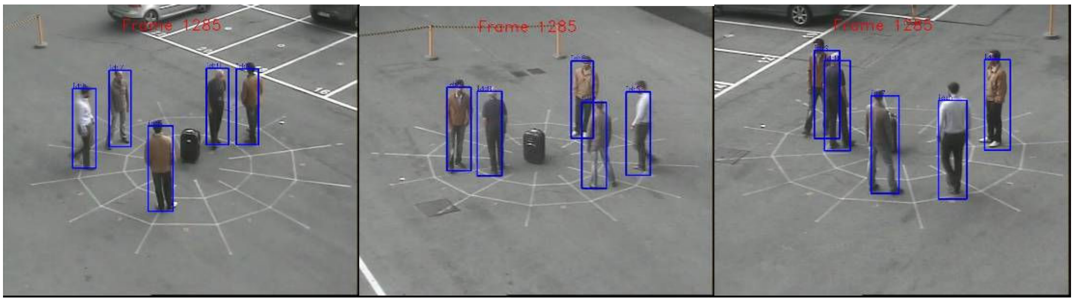


Figure 2-6: Example of person detection in multi-view scene

Images captured simultaneously from three separate cameras surveying the same scene from different angles

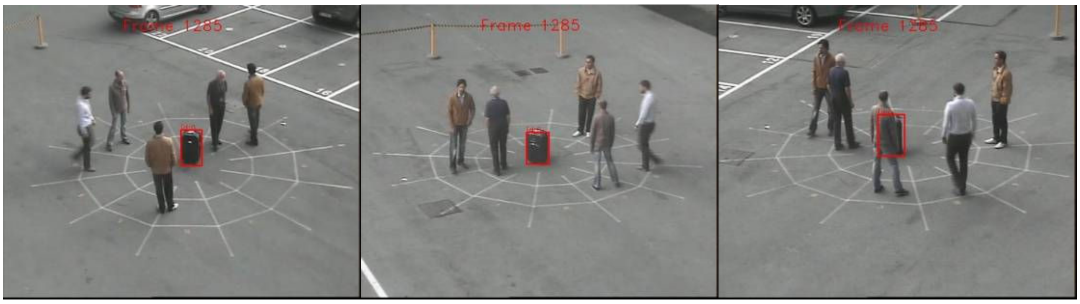


Figure 2-7: Example of bag detection in multi-view scene

As with Figure 2-6, the images shown were captured simultaneously from three separate cameras surveying the same scene from different angles.

The person and baggage detection algorithms developed during the project achieved the aims of providing robust localisation of moving objects alongside objects that have become stationary. Improvements to the state-of-the-art component algorithms, as demonstrated, have been implemented while being able to maintain a relatively fast rate (~6Hz) at which data could be output to the tracking module, [cf. Objective 3]. Further algorithmic improvements utilising Multi View-Infinite plane techniques have increased the speed of operation to between 15 – 30 fps.

2.5.2 Use of other cameras study

While the core object detection algorithmic development focussed on the use of fixed CCTV camera installations, an alternative approach was also considered using PTZ cameras to achieve similar goals of robust people and baggage detection.

PTZ cameras have the capability to focus on any part of the scene at any resolution, although this has the drawback of losing awareness of any event occurring at other locations. In the context of unattended baggage detection, the SUBITO project considered the approach of a “guard tour” where a set of positions for the PTZ device with sufficient resolution to achieve robust object detection was defined to create a mosaic of the surveyed scene. The PTZ camera can be controlled to repetitively cover the defined set of views and as such provide a representation for every part of the scene, the update time being dependant on the length of the tour.

To highlight the capability of this type of system, a PTZ sub-system was implemented as defined in Figure 2-8. A few of the issues relating to this sub-system are discussed below.

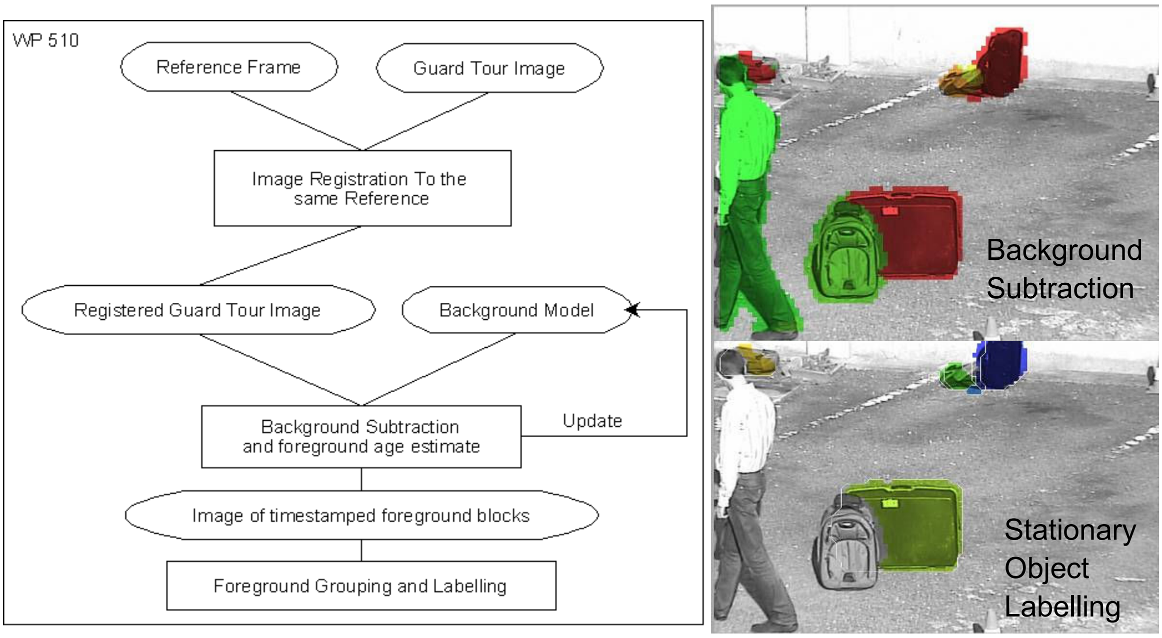


Figure 2-8: PTZ Camera Sub-system

- PTZ Guard Tour Image Registration - The guard tour approach is subject to the mechanical errors within a camera pointing system and as such no two consecutive images of the same guard tour position will be the same. Thus they must be aligned prior to any image processing algorithm.
- Background Subtraction - A sequence of images of a guard tour position is used for background image subtraction. An algorithm was especially developed for this purpose and has been published in [Ref. 23]. Any detected foreground is also used to estimate the age of subsequent image foreground. The resulting image contains the age of any detected foreground block.
- Object identification - The background subtraction output does not include the notion of objects. An additional tool was developed to group the detected foreground ‘blocks’ into stationary objects that are present in the scene for more than a prior defined time *t*, in this case 60 seconds. This algorithm is based on age homogeneity with the capability to deal with object occlusion.

Further details of the PTZ object detection can be found in Refs. 7 & 13.

2.5.3 Object Tracking

The output from the robust object detection is passed to a tracking module in order to perform robust and long term tracking, [cf. Objective 1]. Here the primary aim was to maintain robust tracks for all the individuals in the scene throughout the sequence in order to provide input into the higher level analysis such as threat analysis.

Robust tracking refers to the real-time generation of reliable “tracklets”, while the long-term tracking refers to the processing of these tracklets to form complete tracks. The tracking module had two main roles: to produce a spatio-temporal detection of relevant entities – a track – and assign a unique ID tag to each track. As the track detection and ID track maintenance functionalities are closely intertwined, these aspects were jointly designed.

Initial design of the tracking module used simple reference algorithms based on linking detected entities through simple background subtraction techniques. Further development followed that focussed on more sophisticated track construction involving multiple hypothesis management utilising the multiple camera detection described above, section 2.5.1. The final development tasks looked at the design of functions for long term track association.

As the primary output from SUBITO was a customer demonstration, it was felt important that the system be able to run as near real-time as possible. Current state-of-the-art tracking algorithms, involving sophisticated detection steps, hardly reach a computational rate of a few HZ. The approach adopted for SUBITO was to initially define a baseline running processing chain that would operate real-time and to which performance improvements could be made incrementally, as follows:

- The baseline approach was a simple scheme using Kalman filter models and Munkres associations to provide tracking for detections obtained using simple background subtraction techniques.
- This baseline was then modified to provide enhanced localisation in space for tracked objects by exploiting multi-view detections to improve the eventual threat detection and reduce the false alarm rate.
- The implemented multi-view algorithm used homographic constraints and *a priori* hypotheses to produce an estimate of the 3D bounding box containing the moving individuals.
- Extra spatial reasoning, using accumulation maps, was implemented to filter out some of the artefacts generated in fusing the data from the multiple views, which proved especially fruitful when the scene contained multiple people interacting, Figure 2-9. More details can be found in Ref. 13
- Further performance improvement was produced by introducing better track and ID management by maintaining several concurrent track hypotheses in a delayed decision making framework based on the Multiple Hypothesis Tracker (MHT). This scheme better manages small occlusions and ambiguities with a minor computation cost when the number of targets and false detections are low.

The implemented multi-view tracking module was fully tested on a number of video sequences. This included a comparison of false alarm rate and performance with existing tracker technologies, (i.e. optimal instantaneous assignment, global nearest neighbour, linear Kalman filter etc.)

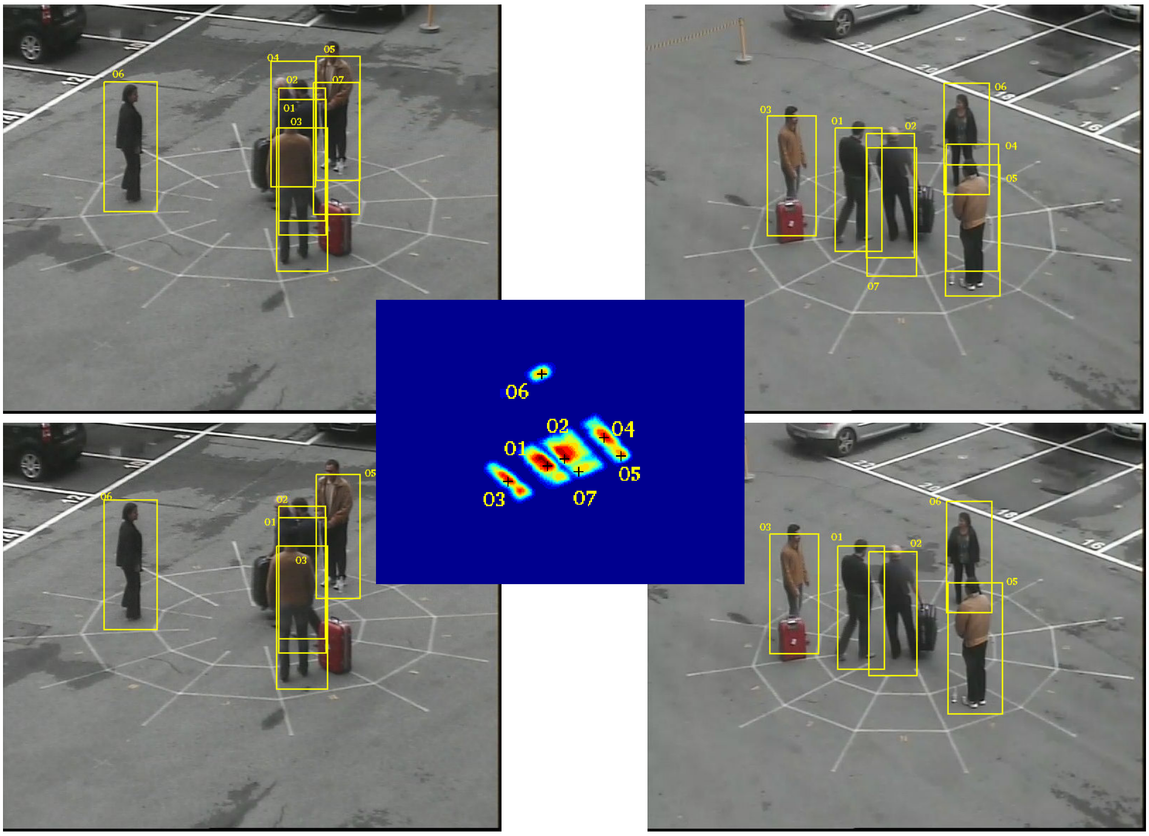


Figure 2-9: Artefact removal using Spatial Reasoning

Example of spatial reasoning on the accumulation map (centre) to filter out false positives (maxima numbered 4 and 7). Upper row: detections from the original map. Lower row: detections after filtering.

A comprehensive set of metrics, (including track recall, precision and latency), was used to gauge the tracking module performance especially in the capability to maintain a unique ID for every individual in the scene throughout the sequence, which is important for identifying the owner of unattended baggage and determining their path, [cf. Objective 1 and Objective 3]. The evaluation of the tracking module results are summarised below. Full details of the evaluation and discussion of the results can be found in Ref. 13.

The quantitative evaluation of the implemented method on a multi-view sequence with ground truth[2], showed a reduction in both the number of false alarms (55%) and of fragmented tracks (46%) compared to the baseline solution, however this is done with reduced mean recall (or track detection rate), most likely caused by the delayed decision-making mechanism of the MHT.

[2] PETS2006 sequence

The implemented tracking module for SUBITO has demonstrated that multi-view tracking in real time is achievable and is able to handle a certain amount of nuisance phenomena (e.g. short term occlusions and ambiguities). It also improves greatly on the tracking performance achieved in the ISCAPS project, although the track ID maintenance on a long term basis in scenes with many interacting people remains a very challenging task.

In general current state-of-the-art multi-object tracking methods can be divided into two groups: those with good performance but which are computationally intensive and real-time methods with a reduced performance. The implemented SUBITO tracking algorithm can be viewed as having achieved a mid-point between the two, achieving satisfactory results at video rate on a standard CPU.

2.5.4 Object Classification

Having robustly detected and tracked objects of interest within the scene the SUBITO system design attempts to perform object classification. This was achieved in two ways:

- Facial recognition
- Object description by use of PTZ cameras.

2.5.4.1 Facial recognition

With facial recognition, the overall performance depends heavily on the quality of the input image, in particular the image resolution of the facial area within the image, typically measured in the number of pixels of inter eye distance (IED), contrast or colour depth within the facial area, sharpness of the objects within the image and pose of the facial image.

The system architecture design work showed, (Section 2.3 and Ref. 5), that standard CCTV would not be sufficient to give high enough resolution for facial resolution algorithms to be used in the SUBITO scenarios and thus the use of a Narrow Field Of View (NFoV) PTZ camera to collect high resolution image was defined, Figure 2-4.

Image resolution is one of the most important limiting factors with standard video-surveillance cameras. In most cases, low resolution prevents object detection at a distance exceeding 15 meters, and prevents face identification in nearly all situations.

The work carried out in the SUBITO project has shown that face recognition can be successfully achieved only in cases of moderate poses (< 25° from head on) with inter-eye distance larger than 20 pixels. Consequently, cameras dedicated to face recognition should be located in a way that can capture channelled person flows (e. g. stairs, passages, entrance and exit points, etc.)

Low IED values normally result in a significant degradation of the face recognition accuracy which would be unacceptable for a SUBITO system. Thus the project developed a concept that would improve recognition performance even in the presence of such variations of IED.

The basic concept implemented was to incrementally improve the facial representation of a person by adding in further sample images of the same individual into the search image or “probe” of the surveillance data.

For example, starting with a single image of a person (probe) after an unattended baggage event was declared, the system will extract faces from the video streams of the remaining cameras and sort them according to decreasing similarity of the person to be searched for, as illustrated in Figure 2-10. In the case of high similarity values the system will (automatically or with the support of an operator) add corresponding images to the set of probe images. After this, the search may be repeated or continued with the enlarged set of probe images leading to possible further hits, Figure 2-11.

Figure 2-10: Facial Detection using single probe image
Top: Single search image; Middle & Bottom: Search results ordered in decreasing similarity values

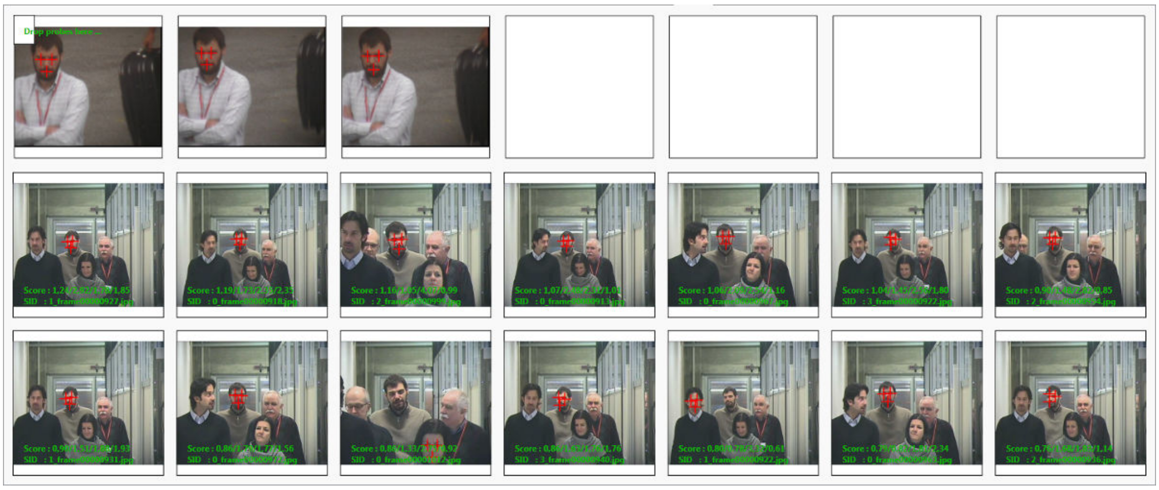


Figure 2-11: Facial Detection using multiple probe images
Top: Probe using three images. Middle & Bottom: System gets further new hits of search individual

In this way the system can iteratively improve the representation of a person covering more and more variations of that individual and thus increasing the probability of recognition at a different camera location and thus to track it across a camera network.

Another key development to come from the work carried out in the SUBITO project was the improvement of the accuracy of the face recognition in the context of the scenario where a person is tracked / recognised across a camera network. In this scenario, the representation of a person differs significantly from camera to camera due to different head poses, varying illumination, camera characteristics, facial resolutions etc.

2.5.4.2 PTZ object classification

The second area of object classification to be considered was object description by use of PTZ cameras. In this case the work carried out in the project looked at the capabilities of a PTZ camera network to provide both stationary object detection and faces associated with long term tracking.

While a PTZ guard tour, section 2.5.2, can provide high resolution images of any location in the surveyed scene; tracking can only be performed effectively and globally using fixed wide angle cameras which provide a constant view of the entire scene. Therefore the work looked at extracting higher resolution images of objects tracked in the wide angle camera network.

The SUBITO tracking module provides tracks for individuals and baggage in the form of a set of 3D trajectories expressed in ground plane coordinates. Given the full camera network calibration defined during the project, [Refs. 2, 7 & 13], these trajectories were intersected with the guard tour positions to extract higher resolution image patches and thus provided additional images of individuals faces and pieces of baggage, Figure 2-12.

Figure 2-12: Image Patch Extraction from PTZ Cameras

Left: detected tracks in fixed camera. Middle: corresponding boxes in PTZ camera. Right: Extracted patch sample.

This concept was extended further, by using two PTZ cameras simultaneously performing the guard tour strategy and each running the single view stationary object detection, a multi-camera/view object matching algorithm was developed. This algorithm extracts the full potential of the epipolar geometry that can be built from using a pair of PTZ cameras allowing the algorithm to deal with objects which are partially or fully occluded in one of the two views, as illustrated in Figure 2-13. Objects that have been identified by the combination of the two cameras can be used to extract additional information such as 3D location, height and volume of stationary objects.

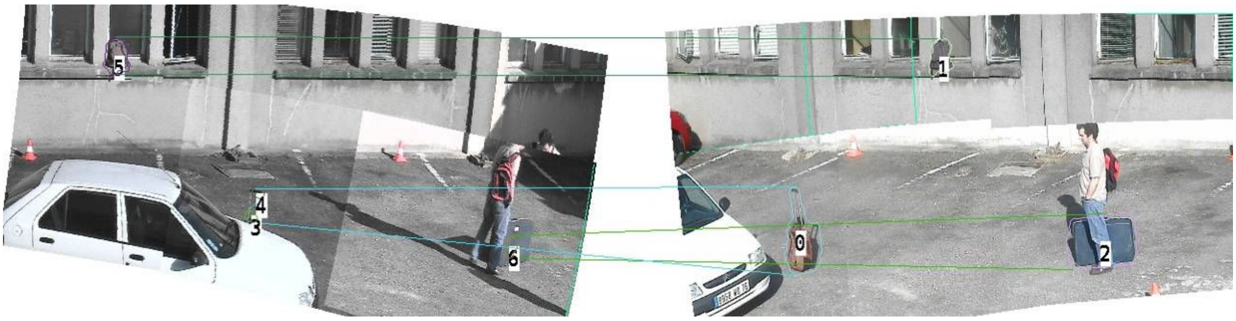


Figure 2-13: Multi-view Object Matching.

Left and right images are mosaics built from 8 guard tour positions. Images here represented are spherically rectified so that matching objects have corresponding heights. The capture is not synchronized: objects may be present in one camera but not visible in the other one.

The developed algorithm was compared to other state-of-the-art methods using several sequences acquired from a guard tour over a 40 x 40 metres area, and was found to provide improved performance over the state-of-the-art methods, [Refs. 2 & 13].

2.5.5 Observe & analyse

The above discussion focussed on the image analysis portion of the system which was designed to perform object detection, tracking and classification. The output from this was fed into the threat assessment sub system, which was designed to analyse the situation and provide a threat assessment of the surveyed scene.

A key objective of SUBITO is to develop a more reliable method for detecting abandoned baggage, [c.f. Objective 1], and reducing the number and impact of false alarms, [c.f. Objective 2]. This was achieved by focussing on understanding the intentions and interactions of individuals within a scene and thereby demonstrating performance that goes well beyond that achieved in the ISCAPS project.

The Threat Assessment work has advanced the state-of-the-art in three principle ways:

- Firstly a method for inferring social relationships between the people in a scene was developed that outperforms existing methods;
- Secondly a method for inferring where individuals are heading within a scene has been developed; and
- Finally a method for detecting abnormal pedestrian movements was developed which compares predicted and actual paths given knowledge of local goals and points of entry and exit.

2.5.5.1 Inferring Social relationships

Social groups are a very common phenomena in human crowds, with empirical studies suggesting that about 74% of people arrive at a social event as a group, and about 50 – 70% (depending on the environment) belonging to a group during casual walking. However despite this fact, crowd behaviour models in current simulation tools, computer graphics applications and in particular activity recognition and computer vision are based on modelling individuals independently.

The work carried out during the SUBITO project has developed a real-time algorithm for automatic detection of social groups within crowds. The method is based on the Social Force Model (SFM) widely used in crowd simulation community. The SFM assumes that the individual's movement is influenced by notional forces, depending on whether the individuals know or do not know each other and correspondingly produces a different set of trajectories in each case.

The method developed during SUBITO solves the inverse problem – given the trajectories, what are the social forces, and thus the relations, that caused that behaviour. The details of the method are expanded slightly below and explained in detail in Ref. 11.

The method was evaluated during development on three publically available video sequences of increasing complexity, ETH and HOTEL sequences from the BIWI Walking Pedestrian dataset[3] and the STUDENTS003 sequence[4]. These particular sequences represent both sparse and complex interactions among individuals.

Figure 2-14: Snapshots from Threat Assessment Test Sequences

Left: The HOTEL sequence; Right: STUDENTS003 sequence. A number of trajectories are highlighted to demonstrate the complexity of the sequences. The straight thick lines connecting individuals indicate group links.

The performance of the developed method was compared with two other methods, (Ge *et al*. [Ref. 24] and Pellegrini *et al*. [Ref. 25]), and evaluated the results by comparing the predicted social grouping of pairs of individuals with ground truth. As expected, for the relatively sparse and simple ETH sequence, all three methods produced similarly good results as most of the groups are isolated in space and time. A larger difference of results was observed for the HOTEL and STUDENTS003 sequences which are much more crowded and complex. One alternative method fails, Pellegrini *et al*., as the distance and velocity based measure is not strong enough to distinguish between group and non-group relations and produces false links. The alternative method performs quite well but suffers from a significant level of false positives being generated.

The best results in the complex scenario are achieved by the inverse-SFM method developed during SUBITO, [Ref. 2].

[3] You'll never walk alone: Modelling social behaviour for multi-target tracking., S. Pellegrini *et al.*, *IEEE 12th International Conference on Computer Vision*, pp 261 – 268, 2009.

[4] http://graphics.ucy.ac.cy/downloads/crowds-software-and-data

2.5.5.2 Inferring individuals destination

Having determined the social grouping it is important that the individual's destination also be inferred. When an individual leaves a bag unattended and moves away, their actions may be entirely innocent if they are going to buy a coffee or newspaper, or may constitute a potential threat if they are moving towards an exit. It is therefore important to predict where someone is heading after they have left a bag unattended. The situation is complicated by the possibility that a destination (e.g. an exit) may be reached by an indirect path.

The work carried out during the project developed a novel method for creating a running prediction of the current destination for each individual tracked within the scene, which is capable of coping with indirect paths or random perturbations from an ideal path. The method uses the assumption that individuals within the scene are moving towards given destinations identified within the field of view of the cameras, and that this is achieved by various intermediary waypoints. It is assumed that individuals will attempt to reach each waypoint at a constant speed and heading plus a random perturbation. By comparing the actual motion to this mathematical model, the plausibility of the different target destinations can be evaluated real-time. Details of the method are included in deliverable D400.1, [Ref. 11]. Figure 2-15 shows examples of person trajectories through a scene.

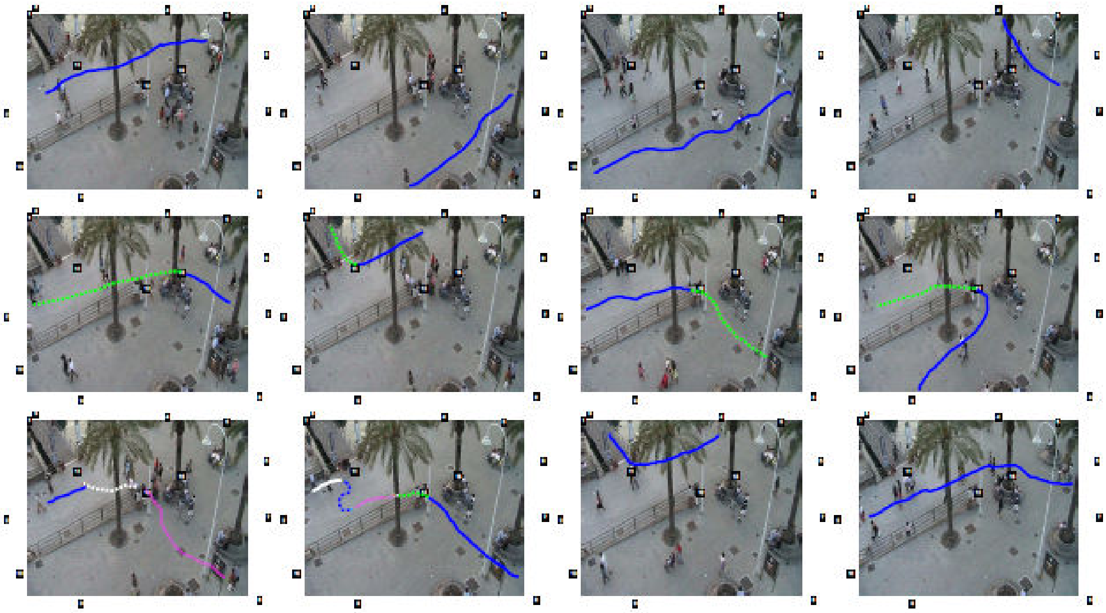


Figure 2-15: Colour-coded trajectories between waypoints, leading to destination

Top Row: Direct route to destination, no waypoints; Middle Row: Route to destination via waypoints; Bottom Row: Errors caused by individuals loitering and not moving according to the assumed model.

2.5.5.3 Detecting abnormal pedestrian movements

In complex scenes, groups of individuals will sometimes have the characteristic of a temporary obstacle, which will cause other individuals to dynamically alter their planned routes to circumvent these groups. The Observe and Analyse subsystem implemented an approach that took into account the current situation presented by the tracking data to create an optimal path prediction. This method for generating situational awareness looks at all available tracking data at any one point in time, and predicts the motion of all the individual tracks for a given increment of time in the future. The tracks are then compared and large deviations provide clues to potentially anomalous situations. Figure 2-16 illustrates the difference between simulated tracks and pedestrian tracks based on an example of the Performance Evaluation of Tracking and Surveillance (PETS) 2006 dataset.

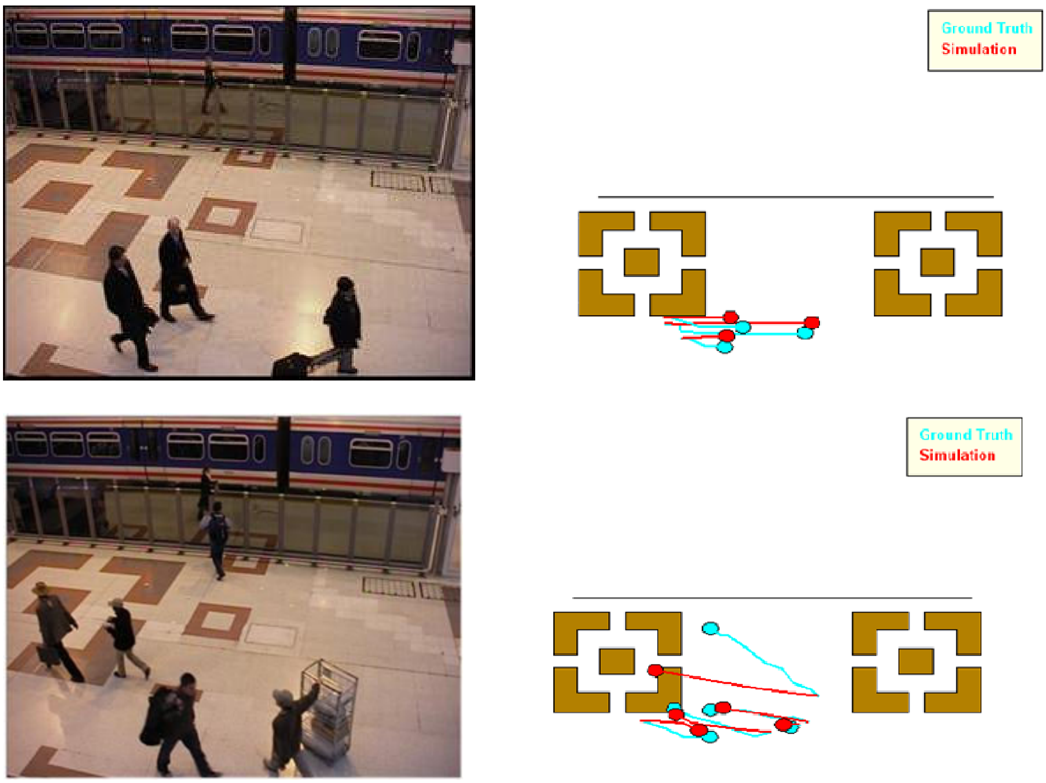


Figure 2-16: Illustration of Predicted versus Tracked Ground Truth Trajectories
Top Row: "Normal Situation" – real trajectory close to simulated path, no anomalous behaviour; Bottom Row: "Anomalous Event" – significant deviation between real and simulated trajectory.

2.5.6 Threat assessment

The methods developed to observe and analyse the situation provided information on the intentions and interactions of individuals within the scene and it was then the aim to develop a threat assessment subsystem to provide a mechanism for reasoning with this information to distinguish between threatening and benign situations.

The design of the threat assessment sub-system focussed around a rule based system which aimed to make it easy to encode the evolving state of the world and explore different behavioural patterns that constitute a potential threat, while maintaining a mechanism that was sufficiently general to accommodate external information (e.g. state of alert, time of day etc.).

The world model is constructed from a set of facts that describe the state of the surveyed scene at the current time-step and pertinent facts from the recent past. In the final

SUBITO design, this model is continually refreshed with the current location of each tracked object and the group structure inferred using the SFM.

The development of the Threat Assessment subsystem was performed in three steps. Each of these steps is described below, and summarised in Table 2-1.

The first "baseline" system was designed to reproduce the performance equivalent to that achieved during the ISCAPS project. The second step added the concept of ownership while the third augmented the ownership concept with social relations.

The baseline system was implemented using the functionality previously explored in ISCAPS, where the threat (i.e. abandonment) was defined by two simple rules:

- Bag unattended if no person within 2 meters
- Bag abandoned if unattended for 30 seconds

The performance of the baseline subsystem was evaluated and the results are shown in Table 2-1. The baseline definition fails to raise the correct alarm in all cases because there is always some individual in the surveyed scene that is sufficiently close to the baggage such that the alarm condition is never reached.

The second step involved implementation of an enhanced threat assessment which included the notion of ownership. In this case the rules were extended to:

- Bag owner is nearest person on appearance
- Bag unattended if owner is not within 2 metres
- Bag abandoned if unattended for 30 seconds

Thus for the threat assessment to be correct, the system was required to raise an alarm following a potential threat in addition to correctly identify both the ID of the abandoned bag and the ID of the owner. Again the results of the system evaluation are shown in Table 2-1. The Ownership definition still makes several errors, but performs better than the simple ISCAPS functionality.

The final step of the development was to augment the notion of ownership by incorporating inferred information on social relations into the threat assessment. This was achieved by running an "inverse-SFM" procedure at each time-step, see Section 2.5.5.1. This enables the system to recognise situations in which someone known to the owner (e.g. a friend or relative) is looking after a bag whilst the owner moves away. Without the knowledge of the social relationship, this situation would otherwise raise a false alarm. The threats are now defined as follows:

- Bag owner is nearest person on appearance
- Bag unattended if nobody from owner's group is within 2 meters
- Bag abandoned if unattended for 30 seconds

Once more the system was evaluated and the results are shown in the final column of Table 2-1. The results using the rules based on Ownership plus Social Groups again show an improvement in performance. Where an error did occur it was through failing to correctly assign two individuals to the same group; which in turn was due to insufficient evidence of their relationship prior to abandonment

Table 2-1: Results for *Ownership* + *Social Groups* Definition.

ELSAG Sequence	Baseline	Ownership	Ownership + Social Groups
18	✓	✓	✓
22	✗	✓	✓
23	✓	✗	✗
24	✗	✓	✓
25	✓	✗	✓
26	✗	✓	✓
27	✗	✓	✓
28	✗	✓	✓
29	✗	✓	✓
30	✗	✓	✓
36	✓	✗	✓
37	✓	✗	✓

The simple reasoning engine used to implement the three versions of the logic outlined above was also used to correct three types of error in the output from image analysis and to improve the assignment of ownership.

- Occasionally changes in the ID of a person or bag due to temporary occlusion occurred, which was corrected by the implementation of a simple rule to detect such situations and substitute the original ID.
- Bags only appear in the scene when deposited by the owner. A bag appearing without an assigned owner was assumed to be spurious and removed from the world model.
- For many scenes, it can be assumed that individuals only appear close to the boundaries of the video frame. The system therefore assumed any person hypothesis that appears away from this boundary was spurious and removed it from the world model.

In the logic implemented for the tests reported above, the owner of a bag is the nearest person to the bag on appearance. Unfortunately, there were cases where the appearance of a bag was delayed due to the nature of the image analysis algorithm, such that a second individual was closer to the bag on appearance. A solution was implemented requiring that the person depositing a bag also paused at the location of its appearance. This additional constraint reduced the chance of an incorrect assignment.

To better understand the actual logic applied by CCTV operators and improve the rules incorporated into future systems, a brief study has been carried out to assess which factors are used to differentiate between benign and threatening situations. The results of the study are documented in Ref. 26.

To visualise the output from the Image Analysis and Threat Assessment algorithms a Man Machine Interface (MMI) was created. The MMI provided a full overview of all available camera images simultaneously and a full resolution output of one source selected by the operator. The highlighting of certain events is introduced by the visualization of the regions of interest which can be switched on or off.

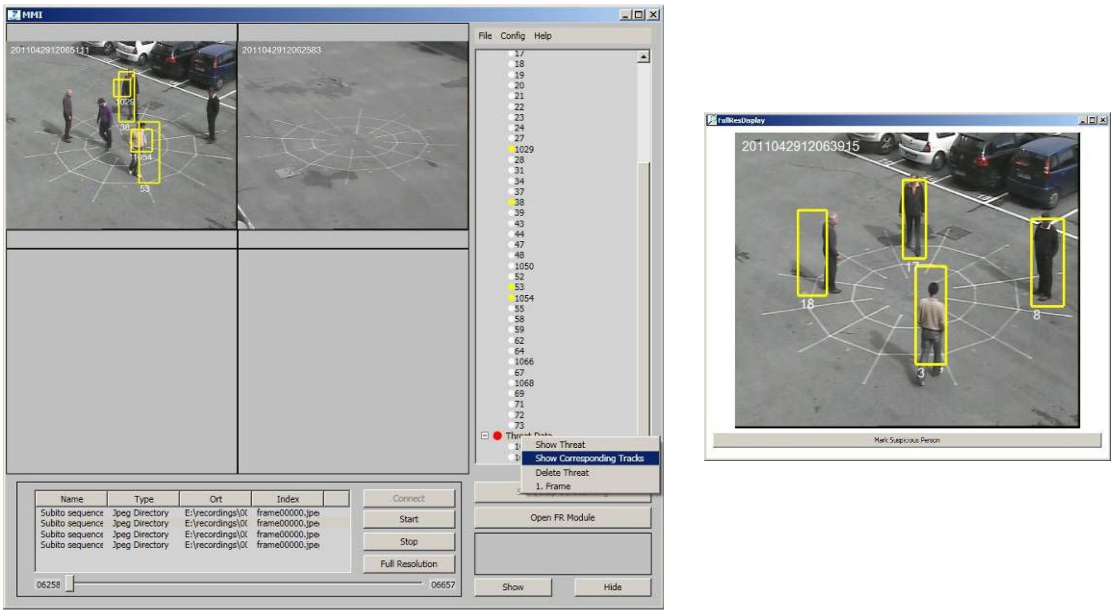


Figure 2-17: Main SUBITO MMI Window & Full resolution camera display

The Graphical User Interface (GUI) is divided into three core areas: the image display area, the camera control area and database control and information area (Figure 2-17, left side).

In the image display area the user has a list of available image sources (live and stored in motion jpeg directories) that can be connected to one of the four output squares of the display areas. One camera can be chosen to display a full resolution image of the scene in a separate window (Figure 2-17, right side).

Furthermore the selected video source is controlled by the video slider providing the functionality of moving backwards and forwards in the video. The database area controls the connection to the database and informs the operator of detected threats and the associated metadata. This metadata is visualized by bounding boxes, lines etc. in the corresponding video field as well as the full resolution display if turned on. The GUI provides all information to identify camera images, with the corresponding automatically generated analysis data like time-stamp position etc. This applies for the recorded data as well.

The MMI is also responsible for the core face recognition capability, i.e. on request of the operator the results of the comparison between two enrolled face images will be displayed and the operator has the opportunity to tweak some of the input data, like setting up the image list associated with one person. Further details on the full functionality of the MMI can be found in Ref. 2.

3 INTEGRATION, TESTING & DEMONSTRATION

The key project objective here was to demonstrate a system capable of automated detection of abandoned goods, fast identification of individual who left them and fast determination of the individual's location or their path, [cf. Objective 3], and to do this the integration process was phased to allow partial results to be incrementally applied.

3.1 Integration

The integration process began by looking at the implementation requirements, drawn from both the SRD, [Ref. 3], and SAD, [Ref. 5], along with discussions with the algorithm development team. Specifically consideration was given to three main areas:

1. Support of the all algorithmic modules operation
2. Support of algorithmic testing
3. Support integration and system test

To support the operation of the algorithmic modules, i.e. system I/O, image analysis, threat detection and man machine interaction, a dataflow architecture was selected. Portioning of the SUBITO data flow was carried out according to the functional needs with the resulting implementation schema shown in Figure 3-1 below.

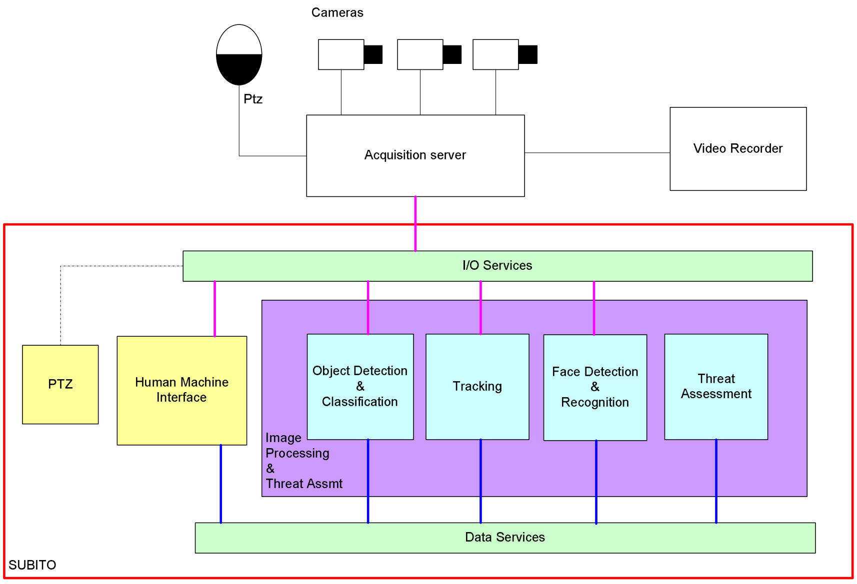


Figure 3-1: Implementation Architecture

The I/O services and collaboration schema was implemented to allow the system to obtain images directly from on-line cameras, video recorded sequences or as a collection of image files. Applications were constructed to a common paradigm which allowed development of the algorithm modules to occur in different environments (Linux and MS Windows) depending on the developers' preference.

The data exchange uses a common database, which guarantees platform independence but pays a price in system performance, although this did not prove to be a constraint in the SUBITO implementation. Only processing results were exchanged via the system database, the video imagery was accessed using a custom API created for that specific purpose.

In general applications implement a simple loop including:

- Read image,
- Process image frame(s),
- Send eXtensible Markup Language (XML) data to the database.

Those applications not requiring images simply loop on the database, reading, processing and writing data therein.

To support the algorithm testing it was decided while defining the implementation requirements that it would be beneficial to utilise the same I/O API at the developers' site to support individual algorithm module testing as was to be used in the final implemented system. Thus the developers' could be confident that the developed modules could correctly access the image streams when integrated with the other component modules.

For support of system integration and test a communication framework was defined which provided a flexible mechanism to switch between online and offline operating modes. This combined approach would be used to verify the system performance allowing recorded test video sequences to be submitted to the system and the resulting output data collected. The same data could also be played back offline, with the MMI allowing the operator to examine in detail the system behaviour and outputs.

With the implementation requirements defined, the focus of the work then moved to the actual hardware integration. The approach adopted was to separate the logical components forming the physical architecture and provide loose coupling of the modules.

The primary requirement was for an efficient yet flexible system, with respect to:

1. Multiple and heterogeneous platform support;
2. Use of field devices;
3. Simplicity and uniformity of the interface;
4. Simplicity of the data exchange mechanism.

The platform support was provided by a collaborative model, which with the exception of image data took the form of XML data, [Ref. 14], exchanged via database tables in a Structured Query Language (SQL) database. Information in the database is synchronised via time tags applied to the data as it is written to the database, which allows the system modules to consume data either one entry (image) at a time or in batches, without the need to know *a priori* the granularity of the data.

The flexibility in the use of test equipment was motivated by the need to be able to incorporate easily changeable configurations, e.g. camera Field of View, orientation etc., however the camera type was limited to analogue cameras as this corresponds to the reality of current surveillance installations and provides an immediate measure of SUBITO with the surveillance market.

The cameras used on the test site were PTZ, but were optimised for orientation and focal length until the most suitable arrangement was found to meet the demonstration requirements, from which point on they were used as fixed cameras. Encoders connected to the cameras translated the analogue signals to digital data for use in the rest of the system

Consideration of the applicability of the SUBITO results to real life situations, e.g. public buildings and spaces, suggested that the test site be of medium to large size, and in practice the area eventually chosen was approximately 40 by 40 metres and can be considered similar to many public buildings.

With the implementation architecture and test site defined, the task of off-line data collection commenced to provide representative test data to the algorithm developers and to provide data for use in the test and validation of the SUBITO results. The scenarios used for the data collection were chosen to satisfy the following requirements:

1. Provide a set of videos covering all the expected project objectives
2. Provide a set of videos containing scenarios that are complete, self explanatory and measurable
3. Perform the recordings according to the standing laws and rules for privacy: personal image rights, privacy of the public, compliance with surveillance cameras laws.

3.2 Testing

Initial testing of the modules was performed using publically available, ground truthed data sets, but these could only test part of the SUBITO objectives. Therefore a set of SUBITO specific scenarios was designed and recorded, [Ref. 16], that would test and validate the SUBITO system against the project objectives. Organised in increasing complexity the scenarios were acted out following precise scripts and constraints. The actors used in the sequences followed exact trajectories that allowed the scenario to be enacted repeatedly but with varying orientations to test system resilience to differing points of view.

All of the defined scenarios contained some basic abandonment action, with some scenarios generating alarms and others only warnings. In each case the basic action flow was:

Individual enters scene, walks to drop point, drops the baggage and either loitered before leaving the scene or left the scene immediately after dropping the baggage.

This simple routine could be played with or without the condition of abandonment being satisfied or used with the alternative warning condition when the individual moved away from the baggage, but remained close enough to or returned before the alarm condition was reached. Scenario complexity was increased by using a second person who accompanied the baggage owner and by introducing additional individuals to the scene both with and without baggage.

The project also considered the issue of privacy in producing the test data sets and implemented a specific procedure for data capture and retention. The procedure imposed certain rules that had to be followed to maintain privacy and data protection requirements:

- Actors should be used when recording the video sequences
- Actors should go through an enrolment procedure that outlines:
    - The scope of the research;
    - The limitations on use of the recorded data;
    - The role of the actor;
    - The period the recorded data will be maintained; and,
    - The person responsible for the safe maintenance on the data.
- A clear definition of limitation of use of the data by other Consortium partners.

When the final integration was completed the test data, as described above, was used to perform a formal test and evaluation of the SUBITO system performance.

3.3 Results

As the implemented SUBITO system was integrated and held as ED's premises in Genova, it was most suitable for them to perform the main testing and evaluation of the final system. To aid in this process SELEX Galileo produced and circulated to the consortium several e-mails and a short technical memo, [Ref. 27], which outlined the requirements and rational for this process, and included suggested procedures for system test, evaluation and data analysis and the way it was intended to report the results in order to address the main project objectives.

The actual testing and evaluation of the final SUBITO demonstration system was performed by ED in October 2011 and recorded in Ref. 28. The following sections (3.3.1 & 3.3.2) provide a summary of the system evaluation. (*Note: The actual evaluation criteria adopted by ED for the testing were not those defined and circulated to the consortium by SELEX Galileo in Ref. 27.*)

3.3.1 Test Results

The report data were obtained by the analysis of the processing results from Leeds threat detection log file, and from observing ONERA track data and the detection results from UoR.

The HMI was not used because it was not essential for the evaluation of the performance results. A tool developed by ED was used to read and interpret the SUBITO data; the evaluation of correspondence with the ground truth has been carried out through manual visual inspection.

Notes on the results:

1. With respect to the format proposed in [2], frame numbers have been introduced to identify the position of the warning and alarm events and for tracking (in this case it is the last frame entry for the owner in the tracker list).

2. Individual notes are sometimes provided by the operator working on visual inspection, to highlight any possible deviation from the expected results and trying to explain the motivation of any ambiguity that may appear in the sequence of results.
3. The success/failure result has been computed by comparing the outputs achieved by all the processing modules in the chain; a successful result is issued all the time that at least one of the modules in the chain has provided a successful output.
4. As multiple additional events can be found for the same instance (these are reported in the outputs from Leeds), they originate multiple output (alert event repetition). These events have been ignored and are not reported in the synthesis tables.
5. To verify the position when the owner is last seen we used the identity of the owner (from threat analysis output).
6. A column has been reserved to refer the output from HMI, which corresponds to what the user operator sees. Through the direct inspection of the video sequence, it is possible to evaluate whether this report is correct or not.
7. In the upper side of each table there is a short description of the test sequence, the number (and identity) of the actors involved in the staged events, date and name of the person in charge of the test, and an image (from one of the camera views) to give a rough idea of the context and the operating conditions of the test

3.3.2 Synthesis of results

The following table refers the results in a table format, for each of the considered sequences.

Table 3-1: Summary of the performance on 15 different sequences

Sequence Nr.	Warning event	Alarm Event	Owner Identification	Tracking	Threat Classification	Success Rate
19	4	0	0	10	0	75%
20	8	8	8	5	8	83%
21	3	0	2	0	1	50%
22	2	0	1	1	0	33%
24	2	0	2	0	0	25%
25	2	0	0	1	0	25%
26	3	3	0	2	3	42%
27	3	0	0	3	1	45%
28	4	4	1	0	4	58%
29	2	0	0	0	0	42%
31	7	0	3	2	2	75%
32	10	9	9	3	9	83%
33	8	9	6	1	4	67%
36	4	4	0	1	4	58%
37	3	0	2	2	1	50%

Very good results, looking at the success rate are obtained in the simple sequences where there is only a person and a baggage (19-20-21-32-33).

Not bad results are obtained in more complex sequences with the interference of other people and baggage (26-27-28-29-31-36-37).

Not good results are obtained in relatively simple sequence with only two persons and one or two baggage (22-24-25).

The success rate is more or less equal respect the previous evaluation with an improvement in complex sequence.

The tracking module works better comparing it to the previous evaluation. There are improvements in the simple sequences and complex ones.

The high level data fusion which leads to owner identification and threat detection is working properly, when the input data (warning and alarm events) are correctly forwarded from the lower levels of the processing chain.

The final score (success rate) is computed as the rate of instances in which at least one of the processing module has issued a correct output value.

3.4 Final Demo

A summary of the SUBITO final demonstration and workshop can be found in Ref. 29.

As part of the final demonstration and workshop, a questionnaire was created to capture the feedback from End Users and interested parties present at the final demonstration, (copies of the original questionnaires can be found in Ref. 2, Appendix B).

The questionnaire first asked the respondents to rate the main goals / capabilities of the SUBITO system in terms of their usefulness/importance in current and future applications. Responses were requested on a rating between 1 (not important) and 5 (very important), Table 3-2.

Table 3-2: End User feedback on SUBITO principle capabilities

SUBITO Goals/capabilities	Current applications	Future applications
1. Abandoned Object Detection	4 (Quite important)	4 (Quite important)
2. Concept of ownership	4 (Quite important)	4 (Quite important)
3. Tracking/locating owner position	5 (Very important)	5 (Very important)
4. Reducing operator workload	4 (Quite important)	4 (Quite important)

Respondents were also asked if there were any other capabilities that they would like in a future SUBITO system. The main replies were:

- Integration of other sensors (thermal cameras, microphones etc.)
- Inclusion of cognitive functions

The questionnaire then asked how well the SUBITO system as a capability demonstrator met the requirements of the End Users. Responses were requested on a rating between 1 (not at all) and 5 (very well), Table 3-3.

Table 3-3: How SUBITO meets End User requirements

How well the demonstrated system meets your requirements for:	Rating
1. Abandoned Object Detection	4 (Mostly)
2. Concept of ownership	3 (Moderately)
3. Tracking/locating owner position	4 (Mostly)
4. Reducing operator workload	3 (Moderately)
5. Functionality, information etc. displayed in the MMI	4 (Mostly)

When asked if there were any aspects of the system which require improvement it was suggested that the general level of complexity of the scenarios, i.e. number of people in the scene, be increased.

In general the feedback for the SUBITO Demonstration system was positive, indicating that the principle capabilities of the system are in line with End Users expectations for current and future systems.

Further work would be required to better meet the End Users requirements, but this would be achieved by inclusion in to the demonstrated system of the more advanced areas of the system architecture, i.e. other sensors (see Section 2.3.1), and algorithm design, Section 2.5.6.

3.5 Competitive Assessment

A competitive assessment task was carried out comparing the SUBITO system against current state of the art unattended baggage technologies. The scope of the assessment was limited to comparing the functionality of SUBITO against existing products in the marketplace. The main constraint of the evaluation was the level of access and detail that could be obtained from sources about the capability of the systems investigated.

It was decided that the assessment would focus on existing products in the marketplace as these are the most mature systems. A number of avenues were explored to identify a list of systems/companies for the assessment. This included: Gathering information from past SUBITO activities, internet searches for products with unattended baggage/object capability, requesting information from partners & in-house capability experts and contacting companies requesting information on the capability of their systems relating to unattended objects.

The assessment focussed on functional capabilities as a full comparison of the functional and performance capability of systems was not appropriate or feasible due to the level of access and detail available.

The criteria used for the assessment was based on the SUBITO primary objectives [Ref 1]:

- Detect UB events
- Identify an owner
- Locate current position of owner
- Pre and post event owner track data for UB event

The criteria also included the basic functionality required (detection, classification, tracking, and behavioural analysis) as well as underlying enabling technologies (3D detection & tracking, PTZ, carried object detection, face detection & recognition, concept of ownership & social relations) that can be utilised to provide a more robust solution (one of the key drivers of the SUBITO project).

An assessment of the underlying enabling technologies and general capabilities of the surveyed products allowed a top level comparison against the main SUBITO objectives to be carried out. The results are shown below in Table 3-4:

Table 3-4: Top level functionality comparison

	Ipsotek	Indigo vision	ELSAG	Object Video	Ioimage	Iomniscient	Analytic Video Systems	BRS labs	Intuvision	Wavestore	SUBITO
Detect UB events	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Identify an owner	No	No	No	No	No	No	No	No	Yes	No	Yes
Locate current position of owner	No	No	No	No	No	No	No	No	Yes	No	Yes
Pre & post event tracking of owner	No	No	No	No	No	No	No	No	Yes	No	Yes

The results showed that, as expected, all the systems investigated have the basic image and behavioural analysis capability required to detect unattended baggage events. Some systems include more complex image and behavioural analysis technologies (3D detection, abnormal pedestrian flow etc.) to provide a more robust solution. But it is only systems which have a form of concept of ownership that can fulfil the main SUBITO functions.

Only one system other than SUBITO (Intuvision) definitely utilise a 'concept of ownership' between objects. No systems other than SUBITO were found to utilise more complex behavioural analysis such as taking account of social relations.

No systems were found which had achieved Image Library for Intellegent Detection Systems (i-LIDS) approval for abandoned baggage detection. A recommendation from this work is that a European-wide and/or worldwide standard would be useful in order to be able to compare the performance of different video analytic systems on a standard dataset.

In terms of performance, a direct comparison could not be carried out (as information was not available for a common dataset) however it has to be acknowledged that SUBITO is currently a capability demonstrator and therefore is less mature than existing products currently in the marketplace. Nevertheless the assessment shows that overall the SUBITO functional capability equals, and in some areas (concept of ownership &

social relations) exceeds, current state of the art unattended baggage detection technology.

Furthermore the fundamental building blocks of the SUBITO system are applicable to other security & surveillance applications. Therefore there is scope that a SUBITO system could be developed into a more wide ranging application in the future.

3.6 Exploitation plans

The SUBITO consortium members have provided information on their exploitation plans to capitalise on the technological achievements made during the SUBITO project. These plans include using the information and the experience gained during the course of the SUBITO project on future activities and these include a number of Framework 7 programmes listed in Table 3-5. The exploitation plans provided by the consortium members are detailed below

Table 3-5 : Future FP7 Programme Involvement

EU FP7 Project	Objective	SUBITO Members participating
ProtectRail	To provide a viable integrated set of railway security solution, by considering:	ELSAG, CEA
Secure-ED	To provide a set of tools to improve urban transport security.	ELSAG, CEA, VTT
EFFISEC	To deliver to border authorities more efficient technological equipment: providing higher security level of identity and luggage control of pedestrians and passengers inside vehicles, at land and maritime checkpoints, while maintaining or improving the flow of people crossing borders, and improving work conditions of border inspectors, with more powerful capabilities, less repetitive tasks, and more ergonomic equipment	ELSAG, Reading U
Co-Friend	To design a framework for understanding human activities in real environments, through an artificial cognitive vision system, identifying objects and events, and extracting sense from scene observation	Leeds U, Reading U
COGNITO	To capture, analyse, store and render with help of 3D graphics complex industrial manual tasks.	Leeds U
VOX-Pol	To create a sustainable critical mass of innovative activity among what is currently a burgeoning, but fragmented group of researchers and research topics with a view to ensuring that EU and Member State strategies and policies targeting VOPE are based on concrete evidence, experience, and knowledge about the contours and workings of VOPE thus increasing their likelihood of success.	Oxford U
FastPass (being evaluated)	To establish and demonstrate a harmonized, modular approach for Automated Border Control (ABC) gates	AIT, Oxford U

SELEX Galileo Ltd.

SUBITO was a successful concept demonstration programme which will aid risk-reduction in the technology field for homeland security. In collaboration with our consortium partners, the results of this concept demonstrator indicated improved performance from that achieved in the previous ISCAPS study with particular improvements in the areas of image analysis and threat assessment algorithms and also in the system architecture by proposing the use of improved camera technology, additional sensors and distributed processing schemes.

Further research and development activity is required to further develop the algorithms and take this type of technology to a readiness level suitable for commercial exploitation; however, the SUBITO project has successfully shown that improvements in these areas have been made and has put forward recommendations where future improvements can continue to be made. A number of the SUBITO partners involved in the detailed technical aspects of the project are planning to work further in these areas and have fostered good working relationships during the project which will be of benefit to future projects.

SELEX Galileo Ltd. activities were limited to management, co-ordination efforts and engineering support and as such did not generate any foreground intellectual property on the SUBITO Project. SELEX Galileo Ltd. Will take forward the invaluable experience and lessons learnt in the role of an EU FP7 Project co-ordinator and intend to apply these to any future EU FP7 projects that they are involved in.

SUBITO has given SELEX Galileo Ltd. an opportunity to better understand the nature of the civil security market and the business drivers and constraints. In particular SUBITO has helped SELEX Galileo Ltd. to better understand the architecture and capabilities of security systems within 'installations' and whether they might be complemented by more advanced sensors, particularly airborne sensors, for more extensive security systems that may extend over many hundreds of square kilometres ('wide zones'). Moreover, SUBITO has aided SELEX Galileo Ltd. to better understand how airborne and ground based security systems may co-operate to produce a more holistic approach to 'pattern of life' based alerts, especially in highly complex urban environments.

ELSAG Datamat S.p.A.

ED as a company is directly involved in the development and delivery of Video Surveillance systems in different fields, the most important application areas being:

- Events security monitoring and control, with the deployment and management of large installation systems (few hundreds cameras) to control important international events like sport Olympics (Torino 2006) or political summits (like G8 2009 in l'Aquila, Italy), Exhibitions and sport events. Video recording and video Analytics are commonly provided to support security operators in their surveillance task, and crime prevention activities.
- Transportations (Railways and Airport). The main reference customer is the Italian Railway company RFI (Rete Ferroviaria Italiana) and most applications are developed in collaboration with Ansaldo spa, (Finmeccanica group). Typical

examples of applications are intruder access detection to galleries, with train wagon identification and classification, as well as security control in the railway stations. Another recent application has been face detection and masking for privacy protection from recorded and displayed videos in the control of public sites in railway stations

- Traffic control, for Municipalities and Highway institutions, law enforcement and land security operations. ED has a long experience in homeland security with specific technologies like mobile and fixed ANPR (Automatic Number Plate Recognition) and hand-held devices for police patrol surveillance and monitoring, as well as sophisticated tools for central operative stations (intelligence operations and investigation procedures).

ED is mainly a system integrator company. As such, the main objective is the satisfaction of its customer with dedicated solutions able to fully solve their problems. In the security and surveillance area a basic component is the integration environment for central supervision and control as well for an effective management of video (both live and recorded streams). A video security integration environment, named S3I, has been developed by ED, to allow an effective integration of all the available commercial cameras and devices as well as special processing tools and Video Analytics functions (both proprietary and third-party solutions). This existing solution represents a reference basis for the SUBITO team, to measure the most significant improvements of the research over the state of the art.

As far as advanced sensory components, the ED lab has developed some specific solutions that are particularly suitable in the video surveillance field:

- 360 panoramic sensor that is able to collect video information from a wide field of view
- Smart cameras (with embedded processing power) that allow video collection at high frequency rate from megapixel cameras (both infrared and color images), specifically used for ANPR applications
- Hand-held device, with a high-resolution camera, to provide many basic biometric control functions (fingerprint, face and license plate recognition)

In the field of Video Analytics a general image processing environment has been developed by ED, with the name of SuperVision [Ref. 30]. This software environment includes all typical basic functions of video analysis, like background-foreground detection, motion detection, motion tracking, and some dedicated functions like face detection and face masking, vehicle detection and classification, etc.

Finally, a fully integrated and dedicated solution has been developed by ED in the area of Homeland security for the control and coordination of all investigative actions dealing with license plate information. The solution, named Auto-Detector in the Italian and European market, and MPH-900 (Mobile-Plate-Hunter) in the US market, represents an essential tool of most police department in the world for an effective control of the territory.

The SUBITO project has a main objective to improve the autonomous detection of alarm events related to the baggage abandonment, identification of the owner and its position.

Among the different technology objectives ED is particularly interested to improved results for:

- Multi-camera baggage identification. The possibility of making detection of abandoned objects from different overlapping images, to make the detection process more reliable.
- Precision tracking of individuals to increase the security system reaction and intervention capability.
- Use of face recognition technology to appropriately pinpoint people positions and if necessary to evaluate further the nature of the alarm by identifying the individual(s) involved.

The contribution of ED team to the exploitation of SUBITO project will be performed along the following lines:

- Identification and description of specific application scenarios, corresponding to the most relevant opportunities in the market.
- Definition of system evaluation procedures to compare the new proposed approach against the quickly evolving innovative solutions emerging from the international market and the products proposed by the most advanced competition
- Integration of some specific subsystem and technical components into the ED proprietary video surveillance platform to demonstrate some relevant technology progress over the state of the art
- Identification of some specific collaborative customer (Railway or Airport organizations) to possibly arrange pilot projects for testing some advanced processing functions as emerged from the SUBITO project

The exploitation planning by ED can be classified in two steps, according to the time frame of the planned initiatives. A short-term plan will be implemented, starting already during the last months of project development, to make use of all the available results. A long-term plan is also devised, to envisage a careful coordination with the business department, aiming to achieve a significant improvement of the position of the company in the security market.

1. Short-term actions.

The main objective of this first phase is a careful testing of some intermediate results, coming from the project research, to verify the real potential in the security field and possibly to get a feedback to the R&D department involved in product development and integration.

- **Use of detection and owner identification**. The identification of some pilot projects with selected and qualified security customers would represent the ideal benchmark to quantify the improvement achieved by SUBITO technology and give us a better feeling about the maturity of the developed solutions in this very challenging area.
- **Use of recorded test set** The SUBITO test set will be used to benchmark proprietary and commercial products.

2. Long-term actions:

This second phase is more oriented to achieve an effective technology transfer into the main product lines of the company in the Security field. As such this initiative will be carried out closely with the engineering department, with particular attention to address concrete application requirements coming from selected customers.

- **Object and baggage and object detection** systems analysis and performance evaluation in a real environment. The main problems are referred to the false alarm rate connected to the use of state-of-the-art products on the market.
- **Use of face recognition technology,** ED estimates that this technology can be profitably used as a complementary technology in many surveillance tasks not only as an official identity verification. The use of this technology can improve the usability of surveillance systems and their reliability.
- **Exploitation on other EU Projects** We have recently started new collaboration in other European Framework 7 projects ProtectRail, Secure-ED and EFFISEC). We are looking at utilising the results of SUBITO on two of these projects(Secure-ED and ProtectRail). On ProtectRail we are exploring the possibility that the results obtained in SUBITO could be used for a specific WP.

ONERA

The exploitation goals of ONERA are mainly to improve its national and international visibility thanks to the results generated during the course of the SUBITO project, and to adapt the developments to other contexts and situations.

ONERA main contribution was to produce a robust and long term real time tracking module for moving pedestrians from calibrated camera video streams. It has been designed according to several hypotheses which may not all be satisfied in real operational conditions or in other settings or viewing conditions.

The module will be exploited as a baseline reference either for other operational conditions (monocular vision, moving camera) or for other functional architectures (tracking by detection, behaviour based dynamic scene understanding, tracklet structure reasoning). One major envisaged evolution is the adaptation of the module to moving platforms with bounded computational capacities especially aerial platforms such as UAVs.

L-1 Identity Solutions AG

Within SUBITO, the face detection and recognition technology was improved especially w.r.t. to low resolution and low quality images. After termination of the SUBITO project, related algorithms will be evaluated further on different data sets. Depending on the outcome of the evaluation process, there are different options for exploitation. One option is to integrate the algorithmic improvements achieved within SUBITO with the standard face recognition technology of L-1 so that basically all face recognition related products could benefit from these. A different option is to integrate the improved algorithms only with specific products. For example the “Face Examiner

Workstation” that is specialized on the analysis of challenging facial image data for forensic applications could benefit from these improvements. The same holds for the

"Argus System" used for screening applications.

On the implementation side, L1 were able to develop some new ideas with respect to integration of other sensor and algorithm input (especially combining detection results related to person identification) and some new approaches in the context of scalability that they will further exploit.

- **Use of other algorithms**

Detection results in image analysis always come with a specific uncertainty represented by a score value. The goal is to keep the False Acceptance Rate (FAR – here: wrong detections) and the False Rejection Rate (FRR – here: missing detections) as low as possible. Without additional input the only parameters to vary (given a specific algorithm) are defining a threshold for this score or some more specific values like the minimum, maximum of inter eye distance, pose tolerance etc. In the SUBITO project, while we were not able to do a thorough evaluation, it still became evident that combing detection results to an overall assessment has the capacity for much more accurate findings. In future projects L1 will (and have already started) to make use of this. The basic design issues that allow both, the independent flow of the different modules (often from different partners running on different platforms) and the structures of data storage and communication that permit an alignment of the various algorithm with its different real time characteristics have already been addressed within the SUBITO system. Parts of these findings will be used in other projects.

- **Scalability**

Face detection and enrollment are computational expensive algorithms. To get the best results the program has to be adaptable to the hardware in use and the search space has to be reduced to the time frames and spatial areas that promise the most valuable output. Within the field of hardware dependency the principles of multiple detection and enrollment threads that we described in the previous sections have been applied to several projects. Work to extend this adaptability to multi server systems is currently in progress. The approach of a ring buffer that allows falling behind if for example a lot of people are in the scenery and catch up to (hyper-) real time processing has proved to be valuable and will be reused. The scope of directed attention, which is analyzing tracks and most common routes and predicting the position of a person in the following frames is another field that occurred in the SUBITO project and will be addressed.

CEA List

In the SUBITO project, CEA List has focused on the use of PTZ cameras for abandoned object detection.

Results have been published in several international conferences and are currently submitted to a journal in the field of computer vision.

This contribution was also a large part of the work of Constant Guillot who has defended his PhD thesis on this topic on December 2011.

Since November 2011, Pierrick Paillet has started his PhD study in our lab to go on this subject : the goal is to add tracking functions of the owner to be combined with the

SUBITO detection results. A platform combining cameras and PTZ is setting up in our facility to facilitate development and integration.

Our objective is now to transfer this know-how to the THALES company through the framework of the VisionLab.

This structure mixing Thales and CEA staff (15 people) has for objective the development of innovative functions for intelligent videosurveillance.

A new function which exploits PTZ sensors in an automatic analysis process would give an decisive added-value to the THALES products, but this is still challenging from a technical point of view. The results reached in SUBITO are a promising starting point for this integration.

University of Leeds

There are four clear directions for University of Leeds future work that arise out of SUBITO, each of which they have started to explore.

The first is to replace the hand-crafting of rules within SUBITO by automatic learning of the behaviour patterns that characterise critical events. We had already made progress in learning simple events in a related EU project (Co-Friend) given video clips depicting examples of these target events, and optionally a spatio-temporal window within the video delineating examples more precisely. In this, we do not use the predominant approach of classifying spatial-temporal distributions of local image/video features and instead focus on the spatial and temporal relations between the objects depicted within a video. This mirrors the representation used in SUBITO. The challenge is to augment this representation to deal with the additional richness of hand-crafted relational models used in SUBITO, such as the critical metrical constraints based on proximity and time that define an abandonment (e.g. person must be out of range of the bag for a prescribed time). We have started by generalising the qualitative spatial and temporal relations in our current work to become learnt relations, thereby allowing adaptation to critical time delays and spatial displacements. This work is on-going within EU project COGNITO. A greater challenge would be to learn the movement models that characterise social groups.

In SUBITO, the logical inference mechanism we used treats visual observations and hand-crafted rules as certain. A natural extension is to associate uncertainty with both the observations and the rules, leading to the inference of critical events with associated uncertainties. We started to do this for rules using Markov Logic Networks (within project Co-Friend) but have yet to extend this to the observations, although there are probabilistic logics that will do this. The implication would be that two incompatible but plausible hypotheses on who introduced a bag into the scene could be explored side by side - this kind of uncertainty was common for the scenarios in SUBITO. This also leads to quantitative assessment of the degree of threat in any given situation, depending on the uncertainty of the facts and rules used in inferring this assessment.

A third direction arising out of SUBITO is to detect anomalous behaviours instead of pre-defined critical events. This would again be achieved by learning, but now within an unsupervised setting. We have made preliminary progress on this using relational models within a new project (Vigil).

Finally, the combination of a state of the art method for event detection based on classifying feature distributions, with relational models over objects (as in SUBITO) shows promise, since it combines complementary information at different levels of analysis and should therefore improve performance.

University of Reading

The University of Reading are exploiting the work undertaken within the SUBITO project in a number of ways. Firstly, the work has been exploited in a tie-up with BAE Systems for the UK Home Office OSCT initiative Technology Demonstrator 2 project on Aviation Security. In this work, the University integrated the abandoned bag detection system with BAE Systems's Universal Threat Management System (UTMS) to produce a unified system for threat detection in public spaces. The system is able to probe abandoned object detections to further establish the threat parameters and to suggest operator response.

The SUBITO work is also being extended in the EFFISEC project on efficient integrated security checkpoints. Specifically, a master-slave Pan-Tilt-Zoom (PTZ) control assembly is being developed to react and provide operators with close up views of abandoned bag detections generated by the SUBITO system.

Finally, the person detection methodology developed by The University of Reading within SUBITO is being extended in order to be able to robustly track people across a range of surveillance scenes.

Valtion Teknillinen Tutkimuskeskus

VTT as an applied research organization is mainly transferring the technology developed during the SUBITO project on a case by case to their industrial customers, e.g. R&D contracts with Sandvik Tamrock, Kalmar Industries, and KoneCranes to develop methods for camera based surveillance for safety of people in front of moving machines.

In addition VTT have also extended the exploitation to distributed wireless security surveillance, where the goal is mainly to decrease power consumption, also in terms of optimised economy in the image transmission.

AIT

Coupling microscopic pedestrian simulation with vision-based pedestrian tracking has promising applications.

The traditional coupling direction and major focus of the AIT Mobility Department is to incorporate large real word trajectory data sets into an in-house solution of a modular and flexible simulation software framework. Real world trajectory data sets are needed to calibrate the parameters of simulation models and to validate simulation results in order to make them as realistic as possible for different contexts.

The accuracy requirements to such calibration and validation datasets are very high, and currently people tracking results from surveillance cameras are in most cases still too weak for calibration purposes, especially for groups and crowds of people. This is why manual annotation of relatively small datasets is often the only way to obtain data from real world scenarios.

AIT Mobility has recently collected indoor real-world calibration datasets with multiple Microsoft Kinect 3D sensors. The SUBITO project was the AIT Mobility Departments only project so far for coupling in the other direction – enriching visual surveillance with microscopic pedestrian simulation. The interesting scenarios for abnormal behaviour detection involve groups and crowds of people (e.g. forming dynamic obstacles), yet also here the weak performance of currently available automatic technologies severely hampers technology development.

Austrian Institute of Technology was restructured during the SUBITO project, and abnormal behaviour detection in visual surveillance is no longer in the strategic fit of the AIT Mobility Department (but rather the AIT Safety & Security Department). The knowledge and experience gained during the SUBITO project will thus be transferred to the AIT Safety & Security Department in the course of future projects dedicated to coupling computer vision and other sensing technologies with pedestrian simulation.

In September 2012 a common nationally funded project will start with Munich airport MUC as an international application partner, where the AIT Safety & Security department will exploit the information of hundreds of already installed surveillance cameras by semi-automatically obtaining 3D models and camera calibration data. 3D information and camera calibration data will support tracking and counting of people within camera fields of view and tracking of salient people over multiple, possibly non-overlapping cameras. SUBITO knowledge will be exploited to support multi-camera tracking over non-overlapping fields of view by coupling simulations with real-world tracks in real-time and comparing the real and simulated trajectories. Other coupling projects will involve management of large events such as festivals, where pedestrian simulations will complement real-time measurements at neuralgic points to obtain situational awareness.

Fiera di Genova S.p.A.

Fiera di Genova was the End User member of the consortium who was intended to provide the venue for the Final Demonstration and who provided invaluable feedback from the perspective of an End User. Fiera di Genova advised that they had no current plans to directly exploit the technology developed during the SUBITO project although they did state that they may reconsider this position at a later date.

University of Oxford

The University of Oxford has found the experience of assessing and advising the SUBITO project on privacy impacts, privacy-friendly design choices, and data protection legislation a useful case study for ongoing research into the problem of designing systems that aim to respect the privacy of users.

The experience gained from SUBITO has provided useful insight into the challenges surrounding practical application of privacy-by-design principles, the integration of privacy concerns in the development of large-scale research projects, and the role that privacy enhancing technologies and techniques could play in alleviating the risks of surveillance technologies.

This research is intended to form the basis of one or more academic publications in the near future, as well as informing and improving privacy design and privacy impact assessment in future research and commercial projects.

The experience gained in SUBITO will directly feed into two new FP7 Security projects that Oxford will be participating in: VOX-Pol and FastPass.

4 DESCRIPTION OF POTENTIAL IMPACT AND DISSEMINATION

The main legal-ethical issues raised by SUBITO are around privacy. A privacy impact assessment was carried out, (see deliverable 200.4 [Ref. 6]), using the general theoretical and practical aspects of the social and legal impact assessment. Where possible, input has been provided to the technical development undertaken in WP400 and WP500 activities. However, there was a significant time lag between the project's original Social, Legal and Ethical (SLE) expert leaving the UK and new SLE becoming a full project partner. There has therefore been a less-than-ideal linkage between the Privacy Impact Assessment (PIA) and technical development processes.

Research was also carried out into the social, legal and ethical requirements related to automated surveillance monitoring of public spaces and identification of individuals. This included an assessment of approaches which are either illegal (or may become so) or which may be unacceptable to the public. Building on the Privacy Impact Assessment and the report produced by Leeds University on the ethics of automated threat detection (D400.2, [Ref. 12]), this research is documented in deliverable 300.4 [Ref. 10].

The ethical impacts of the approach to threat assessment developed within SUBITO was investigated. To place this in context, a range of different threat assessment systems were considered. Manual assessment was seen to be ethically flawed in a number of ways, not least because of individual prejudices and natural limitations of the human operator. As such, automated threat assessment offers a promising alternative as a means of eliminating the human element to overcome these flaws and eradicate prejudice. However it was seen that prejudice can still enter an automated system and, through its presence in the underlying code will become institutionalised. A second suggestion of automation was the eradication of false positives, something which was seen to be true for the system, but not the end user. Similarly, the concerns of function creep and distance between surveillant and surveilled were seen to at least continue if not exacerbate under automation.

While automated threat assessment does have some ethical advantages over its manual alternative, then, these are counter-balanced by competing problems introduced by the nature of automation. SUBITO, though, is neither manual nor fully automated, but a partially-automated or “advisory” threat assessment tool. It is designed to aid operators in dealing with information overload and recognizing threats which may be missed through operator error. This it achieves by sorting people on the behavioural criteria of leaving baggage unattended. By specifying the precise nature of the threat SUBITO both limits its applicability and its ambiguity in a way that defining threats as “suspicious behaviour" does not. Having identified targets SUBITO applies two further criteria of analysing behaviour in crowds and goal prediction to filter out false positives. These are reduced still further by continued human input ensuring that operators and ground staff are involved in the situation, a factor which also helps to counter dehumanization.

That said, there are still ethical concerns regarding SUBITO, not least the possibility of institutionalizing prejudice through the use of the way in which group relationships and intentions are inferred. In both cases the analysis is based on data sets drawn from European contexts and so may prove to demonstrate a bias in favour of European subjects. This could lead to a number of false positives from a similar socio-ethnic background, leading to their being treated unequally. It might also reduce operator trust in the system, and so to an increase in false negatives, if the alerts always concern similar socio-ethnic groups. Overall, though, the analysis within the SUBITO system serves to reduce the number of false positives.

There is little obvious scope for function creep or abuse of SUBITO as a system. The parts of SUBITO and particularly the potential of the threat analysis to identify minority groups or undesirables, or to instil chilling effects on a population, do have potential for function creep. It is therefore important to recognise that the ethical concerns related to the project go beyond the finished product itself. There is also a concern regarding function creep and automated threat analysis as a concept. As automation is trialled and found to be successful in limited contexts, so it will become more popular and more developed. Unless ethical evaluations are provided at every stage of development, what is seen as acceptable at this level of automation and in this context might not remain so as the technology is taken in new directions. It is therefore essential that this is not seen as the final word on the ethics of SUBITO or on automated assessment in general, but as part of an ongoing dialogue.

A full listing of the dissemination activities carried out by the project is included as part of the standard format for the on-line final report, Ref. 31.

5 PUBLIC WEBSITE AND CONTACT DETAILS

For further information see the project website: www.subito-project.eu, or contact the project coordinator:

Mike McPartlin

Project Manager

SELEX Galileo

Tel: +44 (0) 131 343 5204

Fax: +44 (0) 131 343 8110

E-mail: mike.mcpartlin@selexgalileo.com