Copy No: .....

ADVISOR
Second Evaluation Report
IST-1999-11287 Deliverable R8.3

Classification: Public.

Author : M. Renard [VIGITEC]

Reviewed : S. Velastin [Kingston University]
C. Attwood [Thales Research & Technology]
D. Bastin [Bull]
M. Naylor [Thales Research & Technology]
M. Thonnat [Inria]
S. Maybank [University of Reading]

Date : 24 January 2003

The ADVISOR Consortium consists of:

Thales Research & Technology	Coordinator	United Kingdom
VIGITEC SA	Principal Contractor	Belgium
Bull SA	Principal Contractor	France
INRIA	Principal Contractor	France
University of Reading	Principal Contractor	United Kingdom
Kingston University	Principal Contractor	United Kingdom

DISTRIBUTION LIST

Copy Number

1	European Commission	Christos Pipitsoulis
2	Thales Research & Technology	Michael Naylor
3	VIGITEC	Michel Renard
4	Bull	Dominique Bastin
5	INRIA	Monique Thonnat
6	University of Reading	Stephen Maybank
7	Kingston University	Sergio Velastin

(Distribution applies to Issue 1 onwards.)

Classification: Public

Documents, which are classified as Public, may be distributed outside the Distribution which appears above.

MODIFICATION RECORD

Version: ADVISOR-DOC-036_(DRAFT_A) R8.3.DOC
Date: 30 September 2002
Author: M. Renard
Purpose: To be used as guidelines for Test Bed 2 event

.....

Version: ADVISOR-DOC-036_(DRAFT_B) R8.3.DOC
Date: 10 January 2003
Author: M. Renard
Purpose: Review by partners

.....

Version: ADVISOR-DOC-036_(Issue1) R8.3.DOC
Date: 24 January 2003
Author: M. Renard
Purpose: Official delivery

.....

Version:
Date:
Author:
Purpose:

Executive Summary

This is the Second Evaluation Report for the ADVISOR project. It provides the second intermediate evaluation performed by the end-users associated with the project, i.e. STIB (metro of Brussels) and TMB (metro of Barcelona).

ADVISOR is a project that intends to improve management of public transport networks through the improved use of CCTV cameras. The project will undertake to integrate several developments in picture content detection and analysis with technologies for annotating video images, and presenting them for inspection by CCTV operators when trigger events occur. Metro stations will be used as experimentation sites for test beds and a demonstrator.

The goal of ADVISOR is to assist human operators by automatic selection, recording and annotation of "interesting" images such as "abnormal" crowd and individual's behaviours.

The ADVISOR system will be made available to End-users through an appropriate Human Computer Interface (HCI) that will allow them to activate some functions and also to be alerted by the system as new interesting events or incidents are detected. This HCI will be used to let them evaluate the system three times during the project timescale.

Three main evaluation milestones have been programmed respectively after the completion of Test Bed 1, Test Bed 2 and Final Demonstration prototypes. They are conducted according to the Evaluation Plan described in Deliverable R8.1.

A second intermediate evaluation has been organised to capture the End-user feed-back when the Test Bed 2 set of developments were ready. As mentioned above, this was the second milestone where all the developments available at that point were put together and made available to all the project partners and attendees. The first difference from. the first intermediate evaluation was that, this time, the developments were integrated with each other. The second difference was that it has been possible to evaluate, at least partly, the system quantitatively.

Test Bed 2 was, by its nature, still an incomplete prototype and it was therefore not easy, especially for non-technical people, to differentiate the several software modules running together on many details, i.e. to differentiate the capture, motion detection, tracking, behaviour recognition and archiving modules. Moreover, due to the level of development achieved so far, only a part of the full evaluation process was possible (e.g. it was not possible to evaluate co-operation in a multi-user architecture with only one HCI). Nevertheless, this second intermediate evaluation has been conducted according to the Evaluation Plan.

Reactions to ADVISOR presentation have been very positive both from STIB and TMB. Some of the evaluators involved in the process were not previously aware of ADVISOR. Most of them did not work actively in the project. This makes us quite confident that their opinions and feed-back are reasonably objective and could not be influenced by their own involvement in the work done.

Although the scope of this second evaluation was still limited, it has been useful to refine a number of issues, especially those concerning the Human Computer Interface specification. This will be reflected in producing a second release of Deliverable R5.1.

From the evaluation, we identified a number of actions to improve various aspects of the system. These were mainly to :

- better focus the detection on cameras that monitor sensible places
- modify a number of wordings, messages, commands, controls and messages to improve the interaction between the HCI and the other modules of the system (especially the archiving and the behaviour recognition software)
- implement some key functions already foreseen in the original specifications but not yet implemented in the Test Bed 2 intermediate version (like bookmarking for example)

This demonstrated that, as expected, the second intermediate evaluation also provided feed-back to the core ADVISOR software that will be taken into consideration for the final demonstration.

Table of Contents

List of Figures

List of Tables

List of Forms

1 Document Overview

This is the Second Evaluation Report for the ADVISOR project. It provides the report of the second intermediate evaluation performed by the End-users according to the Evaluation Plan previously defined (Deliverable R8.1 – ref. 5 in §1.1).

We have structured the present document keeping the following objectives in mind :

- establish the link with the contractual tasks as defined and described in the project programme (Technical Annex – ref. 1 in §1.1)
- explain how the outcomes of other project work-packages are exploited

The present deliverable includes :

- Document Overview
  - This section gives an overview of the sections in the document; it includes references and a list of definitions for acronyms.
- Introduction
  - This section gives a brief introduction to ADVISOR Work-package 8 and on the tasks covered by the present deliverable.
- Second Intermediate Evaluation Report
  - This section presents the results of the evaluation performed by the End-users at the Test Bed 2 level of the project.

1.1 References

The following Project specific documents are referenced:

Reference 1.
IST Proposal Number IST-99-11287, Annex 1 – "Description of Work" ADVISOR Project Issue 2 dated 01/05/2002

Reference 2.
IST Proposal Number IST-99-11287 V_1.xls A0 – Contract Preparation Forms (CPF) for the ADVISOR project dated 25/10/99

Reference 3.
Project Presentation IST-1999-11287 Deliverable R0.1dated 03/07/2001

Reference 4.
Operator HCI Specifications IST-1999-11287 Deliverable R5.1 dated 20/09/2001

Reference 5.
Evaluation Plan and First Evaluation Report IST-1999-11287 Deliverables R8.1and R8.2 Issue 2 dated 28/02/2002

1.2 Abbreviations and Definitions

The following Project specific terms, abbreviations and definitions apply:

ADVISOR	-	**A**nnotated **D**igital **V**ideo for **I**ntelligent **S**urveillance and **O**ptimised **R**etrieval (this project)
EC	-	European Commission
IST	-	Information Society Technologies
CONVERGE	-	A project within the 4th Programme that has supported activities of Transport Sector projects in the areas of consensus promotion, system architecture, validation, and standardisation. It has helped in the production of design and application principles for in-vehicle human-machine interfaces, identified key user needs for transport telematics services, and prepared a final synthesis of all transport projects' evaluation results
STIB	-	**S**ociété des **T**ransports **I**ntercommunaux de **B**ruxelles (Brussels metro)
TMB	-	**T**ransports **M**etropolitans de **B**arcelona (Barcelona metro)
WP#	-	Work-package number
HCI	-	Human-Computer Interface
CCTV	-	Closed Circuit Television
TECHNICAL ANNEX	-	IST Proposal Number IST-99-11287, Annex 1 Issue 2 – "Description of Work" ADVISOR Project dated 01/05/2002

2 Introduction

This introduction relates to ADVISOR Work-package 8 and the tasks covered by the present deliverable R8.3 within the overall context of the project.

2.1 ADVISOR Project Overview

ADVISOR is a project that intends to improve management of public transport networks through the improved use of CCTV cameras. The project will undertake to integrate several developments in picture content detection and analysis with technologies for annotating video images and presenting them for inspection by CCTV operators when trigger events occur. Metro stations will be used as experimentation sites for a demonstrator system.

The goal of ADVISOR is to assist human operators by automatic selection, recording and annotation of "interesting" images such as "abnormal" crowd and individual's behaviours.

Since CCTV operators would usually have potentially thousands of cameras available at the same time, but only a limited number of monitors, the assistance provided by ADVISOR should increase their efficiency, and help compensate for limited human attention span. ADVISOR will thus generate better use of transport infrastructure by improved safety and security of the environment.

Whilst there are many potential applications within the public transport sector at large (train, metro or bus stations, airports etc.) and other similar environments where "interesting" crowd or individual's behaviour might be detected (city centres, shopping malls etc.). ADVISOR will be focussed on metro stations.

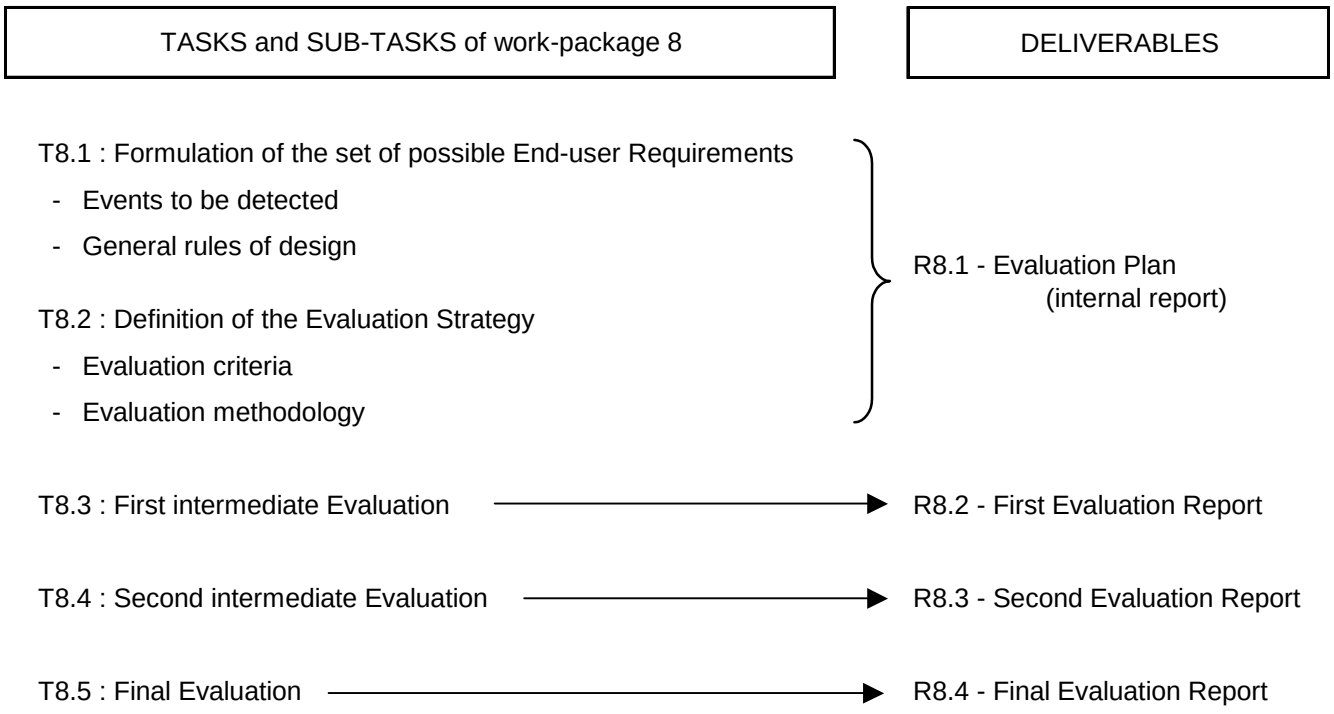
2.2 ADVISOR Work-package 8

Work-package 8 deals with Evaluation and Assessment of the ADVISOR system from the End-user point of view.

As described in the project Technical Annex (see 1.1 above), three main evaluation milestones have been planned, i.e. after the completion of Test Bed 1, Test Bed 2 and Final Demonstration prototypes, respectively. These three evaluation steps are performed in tasks T8.3 to T8.5.

In order to conduct this evaluation process according to the three steps mentioned above on the one hand, and to be able to capture inputs from different persons and organisations on the other, a consistent evaluation plan has been set-up. This plan includes the definition of the end-user requirements that have to be considered for the evaluation, as well as the strategy to be followed during the whole evaluation process. These definition tasks are performed in tasks T8.1 and T8.2.

As usual, the work achieved has to be reported in deliverables. Although the evaluation plan had to be prepared in an internal report (deliverable R8.1) which was not required to be published, we published and submitted it together with the first evaluation report (deliverable R8.2).
The table below summarises and clarifies the relationship between the tasks of work-package 8 and their outcomes.



2.3 Other work-packages linked with Evaluation and Assessment

It is relevant to mention here that other project work-packages in the work achieved have direct impact or relationship with the present work-package 8 on Evaluation and Assessment.

WP1 : End-user Requirements

End-users have participated to describe the security problems they encounter and to define their needs and priorities that finally constitute a list of events or situations that would be processed within the project.

The “Final End-user Requirements” have been described in another Deliverable submitted earlier in the project schedule (R1.1). This outcome of the work package has been used to prepare the Evaluation Plan (R8.1).

WP5 : Human-Computer Interface (HCI)

The ADVISOR system will be made available to End-users through an appropriate Human Computer Interface (HCI) that will allow them to activate some functions and also to be alerted by the system when new interesting events or incidents are detected.

The HCI specifications are described in deliverable R5.1.

In practice, “made available to End-users” means, in the framework of the project , that the ADVISOR Human Computer Interface (HCI) is prepared for and will be used during the intermediate steps of and final evaluation of the system.

WP7 : Validation

The scientific and technical performance of the ADVISOR system will be tested and validated in a separate work-package (WP7). It deals with the measurement of performance and efficiency such as response time, detection rate, false alarm rate, processing speed, etc.

This validation will be performed by the project partners using the same incremental process as the one envisaged for the evaluation by the end-users. It means that test beds 1 and 2 as well as final demonstration will be followed by both a validation process (WP7) and an evaluation process (WP8).

2.4 Project Assessment

Evaluation and assessment is a key step in the development and implementation process of a system such as ADVISOR. In view of further possible industrialisation, any decision

- whether the design or functionality of the system should be modified and updated, or
- whether and how the system should be implemented,

should be made on the basis of sound knowledge about the performance and impacts of the system.

Assessment has been defined as *"the process of determining the performance and/or impacts of a candidate system, usually in comparison to a reference case (existing situation or alternative systems), and usually including an experimental process based on real-life trials, often involving users"*[1].

Firstly, we have to recall that ADVISOR is a research project that aims at realising – in a laboratory - a final demonstrator that will be installed in a real metro station for a few days at the end of the project schedule, i.e. for a short duration. Consequently, the following factors have to be taken into account :

- The installation of a pilot application has not been foreseen in the contract, which means that the "experimental process based on real-life trials" will be replaced by laboratory off-line testing based on pre-recorded video sequences
- There is no reference case that could be utilised (neither an existing nor an alternative system). Even the conventional operations could not be considered as a reference since the local operator has only a reduced number of available monitors on which he displays images "on demand". He has currently no possibility to monitor all the cameras of his station simultaneously.

Secondly, as mentioned above (2.3), the intrinsic technical performance measurement will be achieved within work-package 7 (Validation). The ADVISOR project overall assessment is consequently achieved by both the technical validation performed in work-package 7 and the evaluation process conducted according to the Evaluation Plan.

As far as assessment and measurement of success are concerned, the following table details the

[1] CONVERGE project (see section 1.2) – Deliverable D2.3.1 : Guidebook for Assessment of Transport Telematics Applications – version 3.2 dated September 1998 – by ERTICO (B) and Transport Research Laboratory (UK)

link to be established between the assessment criteria described in the project Technical Annex (reference 1 in §1.1 above) and both the processes of technical validation and evaluation.

Project assessment	Validation	Evaluation Plan ref.[2]
Demonstration of computer vision techniques operating on compressed digital video inputs	X	-
Integration of the techniques via open interfaces	X	-
Demonstration of detection of anomalous events : – detection rate – false alarm rate – response time	X	-
Sustained recording of multiple video inputs in a format that allows efficient retrieval of data	X	3.5.2.2
Demonstration of improved performance in detection and recognition of anomalous events through learning via feed-back from the operator	X	3.5.2.3 3.5.2.4
Quantifiable reduction in operator workload in terms of faster response to incidents and better management of CCTV resources	-	3.5.2.2 3.5.2.4
Increased awareness in the operators of the flow of people through their network	-	3.5.2.1
Use of low cost, commercial technology	-	3.5.4
User acceptance	-	3.5.2
Impact analysis	-	3.5.3
Social cost-benefit analysis	-	3.5.4
Economic analysis	-	3.5.4
Technical analysis	X	-

Table 1. Link between project assessment and validation / evaluation processes

[2] The references mentioned are the sections of Deliverable R8.1-2 Issue 2

3 Second Evaluation Report

The second intermediate evaluation has been organised to capture the End-user feed-back when the Test Bed 2 set of developments were ready.

A one-day meeting was organised on 12/12/2002 at Thales Research premises to present the results obtained so far.

This was in fact the second milestone where all the developments available at that point were put together, although it was the first time where these were integrated with each other, and made available to all the project partners and the End-users present, i.e. STIB (metro of Brussels) and TMB (metro of Barcelona).

The ADVISOR Evaluation Plan defines a number of objectives and a strategy based on sound understanding of the User needs and requirements. It is based on a generic assessment process scheme and identifies six key stages, which are illustrated in the figure below.

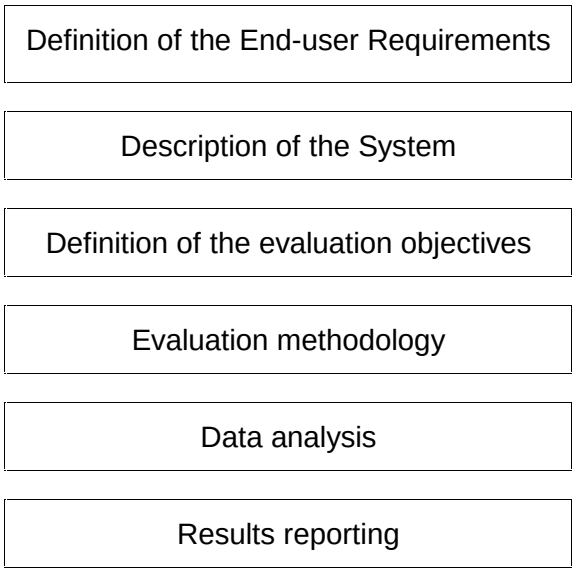


Figure 1. ADVISOR Evaluation Plan

Test Bed 2 was still far from the final complete prototype and consequently, only part of the full evaluation process was possible. Nevertheless, this second intermediate evaluation has been conducted according to the relevant parts of the different stages of the Evaluation Plan described above.

3.1 Definition of the End-user requirements for Test-Bed 2

User requirements have been captured in the framework of WP1 at the beginning of the project.

End-users have expressed functional and operational requirements for the ADVISOR system, i.e.:

(a) what the system has to detect so as to help them to solve some problems encountered, and

(b) how the system has to interact with the operators and security staff in order to be

efficiently integrated into their security management.

These End-users requirements have been described in section 3.3 of the Evaluation Plan. For Test Bed 2, only a subset of them has been addressed at the current stage of the project according to the development achieved so far. In summary, these were :

3.1.1 Required events to be detected for Test-Bed 2

The main effort has been focused on developments able to detect the situations most required by the End-users (recorded in Deliverable R1.1). In accordance with the decisions of the project Steering Committee and taking the results obtained so far into account, it has been decided to use the following events/situations for Test Bed 2 presentation :

Interesting crowd movements

- blocking entry/exit

Individual delinquency

- violence

The evaluators have been in a position to verify how the ADVISOR system should detect such situations.

3.1.2 Required general rules of design for Test-Bed 2

As described in more detail in section 3.3.2 of the Evaluation Plan (see reference 5 in above section 1.1), the complete list of the general rules of design are :

- Designed with the general principle of assisting the operator instead of attempting to replace him
- Managing the alarms (detected events or situations) according to existing rules (information – management – acknowledgement)
- Possible automation or semi-automation of tasks not requiring operator's decision
- Configurable according to company's policies, time schedule, and operator's responsibility
- Designed for integration into an hierarchical architecture
- Secure access to operations and information
- Easy-to-use and adapted to the level of understanding of the users
- Adapted to the working environment

From the above list of requirements, the following ones could be addressed by developments achieved at Test Bed 2 milestone :

- Designed with the general principle of assisting the operator instead of attempting to replace him
- Managing the alarms (detected events or situations) according to existing rules (information

– management – acknowledgement)

- Easy-to-use and adapted to the level of understanding of the users

Mainly, these requirements are addressed through the functionality of the Human Computer Interface.

To be clear, one should say that as far as the system is not able to detect an interesting event and/or to alert the operator accordingly because the necessary software is not yet developed (how it should be at Final Demonstrator stage), it is not relevant to try to evaluate whether the alarm corresponding to this event is rapidly provided or not. For all the other requirements, similar reasoning could explain why they are temporarily out of the evaluation scope.

3.2 Description of the system for Test-Bed 2

Since the system will be evaluated by different people working in different organisations, in different countries and currently operating various security systems, it is necessary to provide them with a clear and concise description of the key characteristics of the system to be evaluated.

3.2.1 General Overview

The ADVISOR system is described in the following documents :

- ADVISOR project work-programme (see §1.1 – reference 1 above)
- Project Presentation - Deliverable R0.1 (see §1.1 – reference 3 above)
- Operator HCI Specifications – Deliverable R5.1 (see §1.1 – reference 4 above)

The two last documents mentioned above have been used to prepare the description of the system for the evaluators.

Test-bed 2 is sited at Thales Research & Technology Ltd. It makes use of data recorded at carefully chosen sites provided by TMB (Barcelona) and STIB (Brussels). It takes inputs from pre-recorded video collected from multiple cameras.

The figure below shows the functional blocks, which are implemented in Test Bed 2.

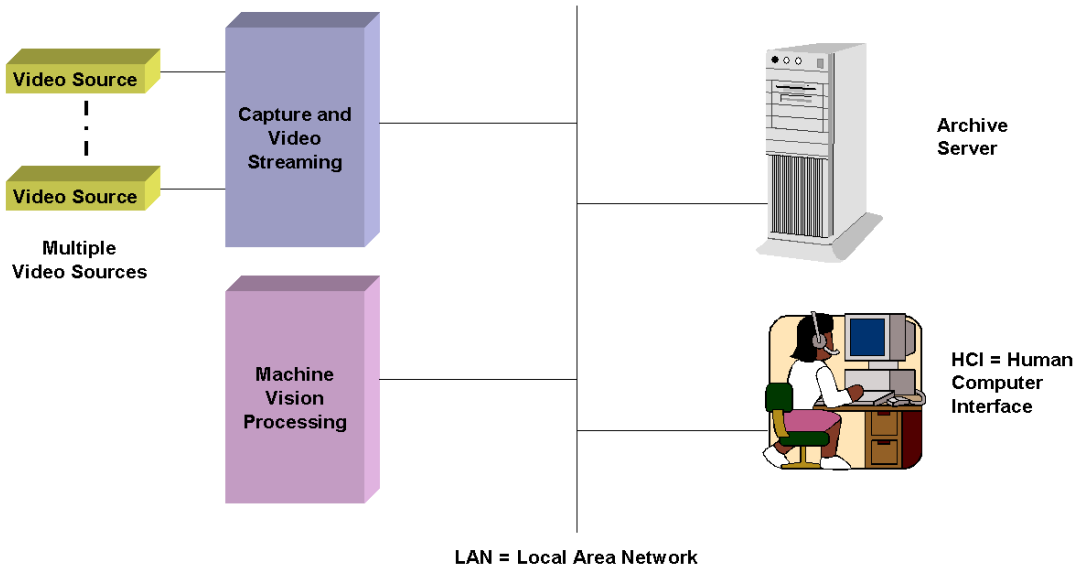


Figure 2. Test Bed 2 system

The ADVISOR Test Bed 2 system has the following four main functions which are described below :

- Capture and Video Streaming
- Machine Vision Processing
- Archive, Search and Retrieval
- Human Computer Interface

3.2.2 Capture and Video Streaming

The Capture module captures and digitises the video input from video sources (interesting sequences recorded at STIB and TMB stations). It compresses the video information to maximise storage capacity. In addition, it adds header information to the captured images (Time stamp). Finally, it transmits the compressed images to other modules within the system.

In addition to operation with external inputs, the capture functionality also includes a mode of operation whereby previously captured video sequences can be played into the system from the hard drive. This mode of operation is useful for system debug and validation.

The ADVISOR system operates with up to four video sources simultaneously. It is capable of operating with either colour or monochrome video sources.

The ADVISOR system captures and processes video images at a nominal rate of five equally spaced images per second per input. Images are reduced to quarter normal size with a square pixel shape and an image matrix of 384 x 288 pixels.

The resultant reduced image is then encoded according to the Baseline-JPEG standard to provide a constant quality image. The target maximum long term mean image record size is 40 Kbytes.

3.2.3 Machine Vision Processing

The machine vision algorithms process the video sources inputs in order to recognise the specific events/situations mentioned above (§3.1.1).

The following scenarios will be recognised by the Test Bed 2 version of the ADVISOR system that will generate an output to be further presented as an alarm to the operator.

- **Blocking or obstructing a recognised entrance or exit**. Blocking corresponds to a situation when a group of people (at least 2 persons) has stopped in a predefined zone for a little while (at least 4 seconds) and can potentially block the path of other people. This situation may reflect a "simple" dangerous situation (talking people block the exit of an escalator for example), but also may be an indication of pick-pocketing activities (the victim seems to be "accidentally" blocked by talking people but in fact it is a way to allow them to steal).
- **Fighting**. Fighting corresponds to a situation when a group of people (at least 2 persons) are pushing, kicking or grasping each other for a little while (at least 2 seconds). Identification of this behaviour will be subjective. It would be clear to a human operator when an incident is fighting. It must be dealt with at any location, at any time and on any day of the week. This situation is probably the one which generates the highest feeling of insecurity.

The following behaviour types would generally be used in the formulation of more complex behaviour recognition. It will however be possible to use any of the following behaviour types to trigger alarms if required.

- An individual or group is stopped
- An individual or group is walking
- An individual or group is running
- A lively group
- An empty scene
- A group in a blocking zone
- A group is stationary
- A group is stationary for a long period of time

3.2.4 Archive, Search and Retrieval (ASR)

The key functions of the Archive, Search and Retrieval process are as follows.

- Storage/Replay of video - Tools to create, maintain and search the archive of annotated images
- Post Incident Analysis - May be performed through the HCI by appropriately querying the archive server.
- Re-transmission of operator selected image sequences over the ADVISOR network
- Support for a single HCI.

The ASR is able to continuously store four video streams at five frame per second. A capacity sufficient for storage of recording at this rate for a period of three days is implemented in Test Bed 2.

At the same time as being capable of continuous recording four video streams at five frames per second, it is possible to retrieve and replay four video streams according to search queries from the Human Computer Interface (HCI) based on the following criteria :

- Type of behaviour (as listed in section 3.2.3)
- Station (one or more)
- Camera (one or more)
- Date/Time (begin/end)

As a result to the queries, the ASR returns to the HCI a list of relevant image sequences available. During the replay of stored images, the following retrieve control commands from HCI will be possible :

- Play forward of an image sequence
- Play backward of an image sequence
- Pause the play-back of an image sequence
- Fast forward of an image sequence
- Fast backward of an image sequence
- Jump to the end of a sequence
- Jump to the beginning of a sequence
- Stop the play-back of an image sequence (in this case, the corresponding screen is switched back to the previously selected video source)

3.2.5 Human Computer Interface (HCI)

The HCI presents the operator with an interface to the system. With the HCI, the operator is able to select live camera views and search the Archive.

The main functionality implemented in the HCI for Test-Bed 2 is:

- **Manual selection of cameras**. This is achieved by using maps and pull-down lists to select the relevant metro station and cameras. The operator is able to define on which part of the quad-split screen he/she would like to see which camera.
- **Ability to simultaneously view the scene from up to four different cameras**. Once the operator has selected up to four cameras on her/his four available sub-screens, s/he is able to double-click on one of the images to enlarge it to full screen mode.
- **Ability to search the archive system**. The operator may decide to search particular image sequences in the Archive system (ASR). To do this, s/he toggles in “search mode” and receives the necessary tools to encode a relevant query (as defined in above section 3.2.4).
- **Ability to retrieve searched images from the ASR**. Once the operator has selected stored sequences available in the ASR, s/he is able to retrieve them and to control the replay using relevant “VCR-like” control commands.
- **Simulated display of most recent alarm messages** in list box on console. In Test-Bed 2, the HCI is not yet able to intercept, decode and display in real time the alarms generated by the Machine Vision Processing. However, to allow the operator to get an idea of the working of the system, off-line generated alarm message will be displayed in the appropriate window of the HCI.

- **Ergonomic and aesthetic design**. This deals with a number of issues like positions of the different windows, buttons, pull-down lists, etc … and also with the size and colour of these elements. More importantly, emphasis has been put to make the sequences of operations logical and to minimize the number of necessary actions by the operator.
- **User-friendly and intuitive according to the familiar Windows paradigm**. This implies mainly the use of the mouse buttons, the different menus available and the information messages provided during operations.
- **Simple, adapted to the skill level of operators**.

3.3 Definition of the evaluation objectives and criteria for Test-Bed 2

The evaluation objectives and criteria are detailed in section 3.5 of the ADVISOR Evaluation Plan (Deliverable R8.1).

In summary, the identification and definition of evaluation objectives primarily needs to be based upon the definition of user needs. What are the key questions to which the users, decision makers and other stakeholders concerned in the project must have answers ?

With evaluation objectives, there should correspond criteria for making judgments and possible choices. The evaluation objectives should relate closely to the implementation and use of the system. At this level of development, one could say that evaluation objectives should address needs and requirements of the security management, the operators and the technical staff of the metro companies. At a later stage (final demonstration), commercial/marketing management as well as public authorities and passengers could be involved.

Test-Bed 2 being essentially an intermediate milestone of technical development, it did not make sense to try to evaluate “impact” or “socio-economic” consequences at this moment. Consequently, only “User Acceptance” evaluation has been performed.

User acceptance evaluation aims to estimate users’ attitudes to and perception of system investigated, usually based on questionnaire surveys, interviews, etc …

As mentioned above, here the users are the operators, the technical staff and the security management of the metro companies associated to the project, i.e. STIB (metro of Brussels) and TMB (metro of Barcelona).

The sub-criteria marked in the table below have been defined as being possible to evaluate at this stage of development of the project.

User acceptance criteria for Test Bed 2

Usefulness criteria	
✓	Relevance of the alarms generated
✓	Sufficiency of the information provided
✓	Worth of the information provided
✓	Work-ability of the information provided
Operating criteria	
✓	Human control
✓	Efficiency of the switching functions
✓	Efficiency of the search and retrieving functions
✓	Alarm management
	Value of added (semi-) automated tasks
Using criteria	
✓	Clearness of information / messages
✓	Adaptation to users' understanding
✓	Easiness to learn
	Adaptation to work environment
✓	General ergonomics of the HCI

Figure 3. User Acceptance Criteria for Test-Bed 2

Some of the criteria marked in the table could not necessarily be fully evaluated at this stage of the project development because some of the corresponding requirements were only partially answered by the Test Bed 2 version of the system (e.g. “adaptation to user’s understanding” have not been “fully” evaluated since the HCI used was not in the mother language of the evaluators) or because some functionality were still not working (e.g. without automatic switching of images according to alarms it is impossible to “fully” evaluate the “alarm management”)

3.4 Evaluation methodology

According to this (fourth) stage of the evaluation plan, we had to determine :

- Who are the evaluators ?
- The evaluation indicators
- The evaluation capture

3.4.1 The evaluators involved

Table 9 of the Evaluation Plan shows which type of evaluators should be involved according to which evaluation criteria.

In order to comply with this rule and taking into account the present evaluation objectives, the following evaluators were invited to the TB2 evaluation :

- Security Operator
- Security Supervisor
- Security Manager
- Traffic Manager

Evaluators of both STIB and TMB were freely selected by the project Responsible persons (upon request by VIGITEC) according to a brief description of the evaluation objectives.

STIB (metro of Brussels) :

Evaluators were :

- Security Manager
- Security Supervisor (also well aware of traffic problems)

TMB (metro of Barcelona)

Evaluators were :

- System Manager (also well aware of daily operator work)
- Engineer (well aware of security problems)

However all the types of evaluators foreseen were not able to attend the Test Bed 2 event, taking into account the evaluation criteria for Test Bed 2, one could say that the evaluators present were able to correctly evaluate the ADVISOR system and to provide the necessary feed-back to the project.

3.4.2 The evaluation indicators

According to the rules defined in the Evaluation Plan, each evaluator had to provide his opinion or feelings on the ADVISOR system so that the results he gave could be further exploited and compared with the corresponding results given by another equivalent evaluator possibly belonging to another metro company.

Therefore, each criteria has been evaluated separately and according to quite strict rules. Evaluators have been requested to provide their judgement by giving a mark from A to D. The meaning of these letters are :

- ✍ A = no, not at all, absolutely not
- ✍ B = less than medium, less than satisfactory
- ✍ C = more than medium, more than satisfactory
- ✍ D = very, very good, very well

User acceptance indicator forms

Below, are given **models** of the evaluation forms that have been filled in by the evaluators.

They had to fill in the forms with their marks, i.e. A to D, according to their evaluation of the criteria. They were helped by the question that "define" briefly each criterion.

Each evaluator has been requested to indicate his job function so that we were able to assess the categories of staff to which he belongs.

The following models have been adapted to Test-Bed 2 evaluation, i.e. that we removed criteria or sub-criteria that had not to be evaluated from the complete models defined in the Evaluation Plan.

Your name : ……..……..	Date of the evaluation : …….			
Please, circle according to your function ---->	Security Operator	Security Supervisor	Security Manager	Traffic Manager
Usefulness criteria \ Event/situation w.r.t. place combination	**Blocking entry / exit**			
	Halls (semi-open area)	Platforms	Corridors	Stairs and escalators
Relevance of the alarms generated Do the alarms (detection) provided by the system have an obvious relationship with my security problems ?				
Sufficiency of the information provided Do the relevant alarms (detection) provided by the system have an unambiguous relationship with my security problem ?				
Worth of the information provided Are the relevant and sufficient information that I receive from the system difficult or impossible to obtain from other detectors ?				
Work-ability of the information provided Do the received relevant, sufficient and worthwhile information allow or help me to take good decisions ?				

Form 1. Evaluation of Usefulness Criteria for Blocking Entry / exit

Your name : ……	Date of the evaluation : ……			
Please, circle according to your function ---->	Security Operator	Security Supervisor	Security Manager	Traffic Manager
Usefulness criteria \ Event/situation w.r.t. place combination	**Fighting**			
	Halls (semi-open area)	Platforms	Corridors	Stairs and escalators
Relevance of the alarms generated				
Do the alarms (detection) provided by the system have an obvious relationship with my security problems ?				
Sufficiency of the information provided				
Do the relevant alarms (detection) provided by the system have an unambiguous relationship with my security problem ?				
Worth of the information provided				
Are the relevant and sufficient information that I receive from the system difficult or impossible to obtain from other detectors ?				
Work-ability of the information provided				
Do the received relevant, sufficient and worthwhile information allow or help me to take good decisions ?				

Form 2. Evaluation of Usefulness Criteria for Fighting

Your name : ……	Date : ……	
Please, circle according to your function ---------->	Security Operator	Security Supervisor
Operating criteria		Quotation
Human control		
Do I receive assistance from the system ?		
In which extend do I keep the control of the system ?		
Do I have the possibility to verify the information provided by the system ?		
Efficiency of the switching functions		
Do the switching functions allow me to perform my job efficiently when no alarm occurs ?		
Do the switching functions allow me to perform my job efficiently when one or several alarms occur ?		
Efficiency of the search and retrieval functions		
Am I always able to introduce all the searching requests that I wish ?		
Do I receive relevant answers to my requests from the system ?		
Am I able to efficiently "travel" within the images made available from the system ?		
Alarm management		
Are the alarm messages easy to exploit ?		
Are the alarms easy to process ?		

Form 3. Evaluation of Operating Criteria

Your name : ……..	Date : …….	
Please, circle according to your function ---------->	Security Operator	Security Supervisor
Using criteria		Quotation
Clearness of information / messages Are all the messages that are provided by the system simple, clear and unambiguous ?		
Adaptation to users' understanding Are all the messages communicated to and actions required from operators well adapted to their level of understanding ?		
Easiness to learn Is it easy to learn and use the system ?		
General ergonomics of the HCI Is the ADVISOR system user-friendly ? Is the ADVISOR system pleasant to use ?		

Form 4. Evaluation of Using Criteria

3.4.3 Evaluation Capture

One of the main problems in such an evaluation is to apply the same process throughout the different possible sessions. A full process has been defined in the Evaluation Plan, however it was anticipated that actual processes followed during intermediate evaluations would be simplified according to available material at those periods of development.

3.4.3.1 Description of the process for Test-Bed 2

In summary, the evaluation capture process has been :

- Installation of the ADVISOR prototype at Thales Research & Technology laboratory in as convenient a situation as possible.
- The project was presented and explained so that the evaluators could understand the context in which they would have to work. In particular, they had to clearly understand that they are seeing a demonstration prototype.
- The prototype was explained and operated by the project partner leading the evaluation session. Questions were answered.
- Afterwards, the evaluation forms and indicators were explained to the evaluators.
- Then, they had the opportunity to use the system. At the end of this phase they should have been able to manipulate the functions they would use in reality without difficulties. Attention was given to ensure that each evaluator worked individually and for about the same duration.
- When the period of working had finished, we requested the evaluators to fill in the questionnaires (forms) shown in above section 3.4.2.
- Each evaluator filled in his questionnaires (forms) individually in a separate quiet room.
- Quick consistency checks were performed and questions asked when necessary.
- Finally, we organised a group discussion with all of the operators who participated and recorded additional comments and clarification.

All the data captured during the evaluation sessions was collected and referenced for further processing.

3.4.3.2 Data gathering protocol

In the next pages, we present the full data gathering protocol according to the process explained above and related to Test-Bed 2.

DATA GATHERING PROTOCOL			TEST BED 2										
Reference step	Operations to be performed \ Participants involved in the process		Project partners	Security Operator	Security Supervisor	Security Manager	Traffic Manager	System Manager	Maintenance Enginee	Marketing Manager	Public Authorities	Passengers
1	Installing the prototype or the mock-up											
1a		*in a lab at project partner's premises*	●									
1b		*in a room at metro company premises*										
1c		*on site at metro company premises*										
2	Meeting the evaluators to explain the context, describe the system and the evaluation process		●	●	●	●	●	●				
3	Demonstration of the prototype by the relevant project partner(s) + answering the questions		●	●	●	●	●	●				
4	Let the evaluator operate the ADVISOR system some minutes with the assistance of the relevant project partner(s) + answering possible practical questions		●	●	●	●	●	●				
5	Let the evaluator operate the ADVISOR system during some hours, using recorded sequences containing events/situations to be detected											
5a		*without any observer*										
5b		*with an observer*										
6	Let the evaluator operate the ADVISOR system 3 to 4 days in a realistic environment using live cameras											
6a		*without any observer*										
6b		*with an observer*										
7	Let the evaluator operate once again the ADVISOR system during some hours, using recorded sequences containing events/situations to be detected											
7a		*without any observer*										
7b		*with an observer*										
8	Meeting the evaluators to explain the forms to be filled in		●	●	●	●	●	●				
9	The evaluators fill in the relevant forms individually											
9a	USER ACCEPTANCE	Form 1 - Evaluation of Usefulness criteria for Rapid increase of crowding level										
9b		Form 2 - Evaluation of Usefulness criteria for Overcrowding										
9c		Form 3 - Evaluation of Usefulness criteria for Unbalanced floor occupation										
9d		Form 4 - Evaluation of Usefulness criteria for Blocking entry / exit		●	●	●	●	●				
9e		Form 5 - Evaluation of Usefulness criteria for Violence		●	●	●	●	●				
9f		Form 6 - Evaluation of Usefulness criteria for Detection of persons without tickets										
9g		Form 7 - Evaluation of Usefulness criteria for Vandalism										
9h		Form 8 - Evaluation of Operating criteria										
9i		Form 9 - Evaluation of Integration criteria										
9j		Form 10 - Evaluation of Using criteria										

DATA GATHERING PROTOCOL			TEST BED 2											
Reference step	Operations to be performed / Participants involved in the process		Project partners	Security Operator	Security Supervisor	Security Manager	Traffic Manager	System Manager	Maintenance Enginee	Marketing Manager	Public Authorities	Passengers
9k	IMPACT	Form 11 - Expected Impact on Safety and Security										
9l		Form 12 - Expected Impact on Traffic Efficiency - Delay										
9m		Form 13 - Expected Impact on Traffic Efficiency - Adaptation of Train Numbers										
9n		Form 14 - Expected Impact on the Number of Passengers										
9o		Form 15 - Expected Impact on the Travelling Hours										
9p		Form 16 - Expected Impact on the Usage of "Dangerous" Stations										
9q	SOCIO-ECONOMICAL	Forms 17 & 18 - Installation costs (equipment)										
9r		Forms 19 to 21 - Installation costs (manpower)										
9s		Form 22 - Training costs (preparation)										
9t		Form 23 - Training costs (manpower of the trainees)										
9u		Form 24 - Operating costs										
9v		Form 25 - Maintenance costs (spares)										
9w		Form 26 - Maintenance costs (manpower for the maintenance of the equipment)										
9x		Form 27 - Maintenance costs (update documentation and further training)										
9y		Form 28 - Summary and calculation of the Implementation Costs										
10	Check anomalities in the answers provided in the forms and preparation request for additional comments from the concerned evaluators		●									
11	Individual systematic interviews with the evaluators about :											
11a	USER ACCEPTANCE	Usefulness										
11b		Operating										
11c		Integration										
11d		Using										
11e	IMPACT	Safety										
11f		Security										
11g		Traffic efficiency										
11h		Personnel motivation										
11i		Passenger behaviour										
11j	SOCIO-ECONOMIC	Estimation of the installation costs										
11k		Estimation of the training costs										
11l		Estimation of the operating costs										
11m		Estimation of the maintenance costs										
12	Additional questions - answers - comments in individual sessions											

Reference step	Project partners	Security Operator	Security Supervisor	Security Manager	Traffic Manager	System Manager

Table 2. Data gathering for Test Bed 2 intermediate evaluation

3.5 Data Analysis for Test-Bed 2

Evaluation captured from both End-user metro companies associated with ADVISOR have been analysed in order to produce relevant and consistent material. The analysed data should then enable possible update and refinement of specifications as well as further development driven by User requirements.

For what concerns its quantitative part, the second intermediate ADVISOR evaluation has produced marks (A to D). To simplify comparisons, those letters have been transformed into figures as follows :

A = 0,05

B = 0,35

C = 0,65

D = 0.95

This very simple system allowed us to process the marks from different operators together.

On what concerns the qualitative part of the evaluation, we first reviewed all the comments made by the evaluators during the individual talks and during the general discussion.

Afterwards, and for each criterion to be evaluated, we tried to extract relevant actions or suggestions to improve the ADVISOR system in view of the final demonstration.

The quantitative tables are presented below in sections 3.6.1.1 - 3.6.2.1 - 3.6.3.1 below.

The additional comments made by the evaluators are presented in sections 3.6.1.2 - 3.6.2.2 - 3.6.3.2 below.

Our review of the evaluation is presented in sections 3.6.1.3 - 3.6.2.3 - 3.6.3.3 below.

Corrective actions are presented in sections 3.6.1.4 - 3.6.2.4 - 3.6.3.4 below.

3.6 Evaluation Results Reporting for Test Bed 2

Below, we review the evaluation forms, comments, suggestions and remarks provided by the operators to each evaluation “User Acceptance” relevant criteria. The complete list of User Acceptance criteria is available at section 3.5.2 of the Evaluation Plan (Deliverable R8.1).

It is important to keep in mind that the End-user evaluators have access to the ADVISOR system only through the Human Computer Interface.

To understand many of the comments or suggestions made, it would be useful to have a knowledge of the Human Computer Interface. Relevant information can be found in deliverable R5.1 “Operator HCI Specifications” (see reference at 1.1).

In the different tables of next sections, the evaluators are named STIB 1, STIB 2, TMB 1 and TMB 2. According to section 3.4.1, this corresponds to the following :

STIB 1 : Security Manager

STIB 2 : Security Supervisor

+/- Traffic Manager (good knowledge of)

TMB 1 : System Engineer

+/- Security Operator (good knowledge of)

TMB 2 : Engineer

3.6.1 Usefulness criteria

3.6.1.1 Table of individual evaluations of usefulness

a) "Blocking entry/exit" situation

Event/situation w.r.t. place combination / Usefulness criteria	**Blocking entry / exit**			
	Halls (semi-open area)			
	STIB 1	STIB 2	TMB 1	TMB 2
Relevance of the alarms generated Do the alarms (detection) provided by the system have an obvious relationship with my security problems ?	B	B	B	D
Sufficiency of the information provided Do the relevant alarms (detection) provided by the system have an unambiguous relationship with my security problem ?	B	B	B	C
Worth of the information provided Are the relevant and sufficient information that I receive from the system difficult or impossible to obtain from other detectors ?	C	D	C	D
Work-ability of the information provided Do the received relevant, sufficient and worthwhile information allow or help me to take good decisions ?	B	C	C	D

Event/situation w.r.t. place combination / Usefulness criteria	**Blocking entry / exit**				
	Halls (semi-open area)				
	STIB 1	STIB 2	TMB 1	TMB 2	Mean
Relevance of the alarms generated Do the alarms (detection) provided by the system have an obvious relationship with my security problems ?	0,35	0,35	0,35	0,95	0,5
Sufficiency of the information provided Do the relevant alarms (detection) provided by the system have an unambiguous relationship with my security problem ?	0,35	0,35	0,35	0,65	0,425
Worth of the information provided Are the relevant and sufficient information that I receive from the system difficult or impossible to obtain from other detectors ?	0,65	0,95	0,65	0,95	0,8
Work-ability of the information provided Do the received relevant, sufficient and worthwhile information allow or help me to take good decisions ?	0,35	0,65	0,65	0,95	0,65

b) "Violence" situation

Usefulness criteria \ Event/situation w.r.t. place combination	**Violence**			
	Halls (semi-open area)			
	STIB 1	STIB 2	TMB 1	TMB 2
Relevance of the alarms generated Do the alarms (detection) provided by the system have an obvious relationship with my security problems ?	D	D	C	D
Sufficiency of the information provided Do the relevant alarms (detection) provided by the system have an unambiguous relationship with my security problem ?	C	D	B	D
Worth of the information provided Are the relevant and sufficient information that I receive from the system difficult or impossible to obtain from other detectors ?	D	D	D	D
Work-ability of the information provided Do the received relevant, sufficient and worthwhile information allow or help me to take good decisions ?	C	C	C	C

Usefulness criteria \ Event/situation w.r.t. place combination	**Violence**				
	Halls (semi-open area)				
	STIB 1	STIB 2	TMB 1	TMB 2	Mean
Relevance of the alarms generated Do the alarms (detection) provided by the system have an obvious relationship with my security problems ?	0,95	0,95	0,65	0,95	0,875
Sufficiency of the information provided Do the relevant alarms (detection) provided by the system have an unambiguous relationship with my security problem ?	0,65	0,95	0,35	0,95	0,725
Worth of the information provided Are the relevant and sufficient information that I receive from the system difficult or impossible to obtain from other detectors ?	0,95	0,95	0,95	0,95	0,95
Work-ability of the information provided Do the received relevant, sufficient and worthwhile information allow or help me to take good decisions ?	0,65	0,65	0,65	0,65	0,65

3.6.1.2 Additional comments of the evaluators on the usefulness of the system

- ***Relevance of the alarms generated***
  - Do the alarms (detection) provided by the system have an obvious relationship with my security problems ?
    - ✍ Difficult to evaluate due to the limited number of sequences demonstrated.
    - ✍ For the blocking situation, probably should it be useful to identify the places where pick-pockets operate according to the “blocking method” to steal passengers.
    - ✍ It should be necessary to limit drastically the false alarms if one wants the system to be accepted. The relevance of the alarms should be “tuned” during a testing period with the real operators and security responsible people.

- ***Sufficiency of the information provided***
  - Do the relevant alarms (detection) provided by the system have an unambiguous relationship with my security problem ?
    - ✍ It should be preferable to replace the wording “alarm” by “alert” to avoid misinterpretation from the operators.
    - ✍ The possibility (offered) to go outside of the sequence boundaries is mandatory from this point of view because sometimes, there is more information before the alarm provided by the system.

- ***Worth of the information provided***
  - Are the relevant and sufficient information that I receive from the system difficult or impossible to obtain from other detectors ?
    - ✍ Yes, it is impossible to obtain that information from other detectors. Only an operator looking permanently to the monitor could detect the same things.

- ***Work-ability of the information provided***
  - Do the received relevant, sufficient and worthwhile information allow or help me to take good decisions ?
    - ✍ It will depend on the operator (degree of autonomy, intelligence, responsibility minded, etc …).
    - ✍ It should be useful to mark the sequence to retrieve it easily in the future.

3.6.1.3 Review of the evaluation results for the usefulness criteria

- ***On relevance of the alarms generated***

  The sequences used in Test Bed 2 for the "blocking entry/exit" scenario appeared insufficiently relevant to three of the evaluators and the information provided did not allow them to understand what was happening.

  The fact that Evaluator TMB 2 gave a good score might be explained by the fact that this evaluator was present when the sequences with the actors were recorded and probably understood this scenario better.

  We should also take into account that the sequences were recorded with actors at night when the station was closed and empty (to avoid any trouble to the normal passengers). Perhaps because of this, quite static people (blocking people) in a nearly empty station did not appeared as a salient problem to the evaluators.

  On the contrary, the score is good for the relevance of the alarms generated in case of "violence" situation (people fighting). Obviously, two people fighting should appear immediately much more salient to any operator.

- ***On sufficiency of the information provided***

  In the case of the "blocking entry/exit" scenario, probably the message is not clear enough and the relationship with their security problem is not obvious.

  On the contrary, for the "violence" situation, the relationship between two people fighting and the security problem of the metro company is nearly trivial. One can understand that the fact that the station was empty has had less impact on the perception of the evaluator facing such a situation.

- ***On worth of the information provided***

  In both situations, the evaluators seemed to recognise that no other known systems could deliver an alarm.

  It appears that, they dislike the wording "alarm" which should not be interpreted as the usual alarms they have in their systems. ADVISOR should say to the operator "please have a look at these interesting images" rather than pretend that "these people are pick-pockets"

- ***On work-ability of the information provided***

  All evaluators agreed to say that reaction time should be short to be efficient in such scenarios. However, according to them, it would depend on the operator himself.

  As far as the relevance of the alarm is not trivial, the risk of non-reaction of the operator becomes rather high.

  It was recognised that since the alarms have been recorded, they could also be exploited off-line afterwards to allow some second level reactions or decisions.

3.6.1.4 Actions and ideas to improve the usefulness of the ADVISOR system

- ***To improve the relevance of the alarms generated***

  For the "blocking entry/exit" situation, it would be necessary to better focus on places where a blocking situation is undoubtedly relevant (for example at the exit of an escalator).

- ***To improve the sufficiency of the information provided***

  One could envisage to refine the wording of the messages provided in case of alarm. For example we could try to replace the message "blocking entry/exit" by a message like "blocking the exit of an escalator".

  The word "alarm" should be replaced by the word "alert" within all the messages provided.

  The commands of the retrieving module could probably be further improved to allow easy verification of the images before and after the alarm period.

- ***To improve the worth of the information provided***

  We see little to do to improve on this criterion.

- ***To improve the work-ability of the information provided***

  It is important to implement the bookmarking function (which is foreseen, but was not ready for Test Bed 2 evaluation).

  Better messages (discussed in the second bullet-point above) should improve operator efficiency.

  Improvement of the archive module control should favour a better off-line exploitation of the system and consequently increase its work-ability.

3.6.2 Operating criteria

3.6.2.1 Table of individual evaluations of operation

Operating criteria	STIB 1	STIB 2	TMB 1	TMB 2
Human control Do I receive assistance from the system ? In which extend do I keep the control of the system ? Do I have the possibility to verify the information provided by the system ?	D	C	C	C
Efficiency of the switching functions Do the switching functions allow me to perform my job efficiently when no alarm occurs ? Do the switching functions allow me to perform my job efficiently when one or several alarms occur ?	C	B	C	D
Efficiency of the search and retrieval functions Am I always able to introduce all the searching requests that I wish ? Do I receive relevant answers to my requests from the system ? Am I able to efficiently "travel" within the images made available from the system ?	C	C	D	C
Alarm management Are the alarm messages easy to exploit ? Are the alarms easy to process ?	C	C	D	B

Operating criteria	STIB 1	STIB 2	TMB 1	TMB 2	Mean
Human control Do I receive assistance from the system ? In which extend do I keep the control of the system ? Do I have the possibility to verify the information provided by the system ?	0,95	0,65	0,65	0,65	0,725
Efficiency of the switching functions Do the switching functions allow me to perform my job efficiently when no alarm occurs ? Do the switching functions allow me to perform my job efficiently when one or several alarms occur ?	0,65	0,35	0,65	0,95	0,65
Efficiency of the search and retrieval functions Am I always able to introduce all the searching requests that I wish ? Do I receive relevant answers to my requests from the system ? Am I able to efficiently "travel" within the images made available from the system ?	0,65	0,65	0,95	0,65	0,725
Alarm management Are the alarm messages easy to exploit ? Are the alarms easy to process ?	0,65	0,65	0,95	0,35	0,65

3.6.2.2 Additional comments of the evaluators on the operation of the system

- ***Human control***
  - Do I receive assistance from the system ?
  - In which extend do I keep the control of the system ?
  - Do I have the possibility to verify the information provided by the system ?

- ✍ General positive opinion

- ***Efficiency of the switching functions***
  - o Do the switching functions allow me to perform my job efficiently when no alarm occurs ?
  - o Do the switching functions allow me to perform my job efficiently when one or several alarms occur ?

    - ✍ Would be useful to also have the images on external monitors
    - ✍ For the selection through maps, cameras should be drawn as much as possible as there are oriented in reality.
    - ✍ Overall map is not very useful, but is however mandatory.
    - ✍ Would like additional selection possibility, using keyboard shortcuts and direct entries (all the stations are identified with a well defined abbreviation and all the cameras are numbered according to rules well-known by the Operators).

- ***Efficiency of the search and retrieving functions***
  - o Am I always able to introduce all the searching requests that I wish ?
  - o Do I receive relevant answers to my requests from the system ?
  - o Am I able to efficiently “travel” within the images made available from the system ?

    - ✍ “Today” and “Now” shortcuts buttons are confusing according to some evaluators and useful according to others
    - ✍ “Cancel” and “OK” buttons usage is not clear.
    - ✍ The query should allow filtering by user annotations
    - ✍ VCR like commands require assistance (especially “pause” and “stop” buttons)
    - ✍ Automatic erasing of images older than a given time is required by the law.
    - ✍ Watermarking would be very valuable to allow use of the images in legal cases.
    - ✍ It should be possible to export retrieved images or sequences in an open format.

- ***Alarm management***
  - o Are the alarm messages easy to exploit ?
  - o Are the alarms easy to process ?

    - ✍ Double lines for each alarm is not necessary

- ✍ An alarm line should not include a reference number that is not useful to the operator
- ✍ An alarm line should begin with date and hour information
- ✍ There are not enough alarm lines visible in the window
- ✍ Reporting facilities should be available. For instance, to get a print-out of all the alarms generated during the month. A function to export registers in a known format (like Excel, Access, Ascii or txt) would be useful.
- ✍ When an alarm occurs and the corresponding image is switched automatically, it should be possible to retrieve its beginning very quickly without being obliged to go into the search mode.

3.6.2.3 Review of the evaluation results for the operating criteria

- ***On the human control***

  As soon as they managed the system, the evaluators had the feeling that they had good control of it and gave good score to that criteria.

  Not a lot of additional comments can be made about this subject.

- ***On the efficiency of the switching functions***

  Depending on the evaluator, they gave good or bad mark to that criterion. What is important is that the Security Supervisor gave the lowest score to this feature, which has certainly to be taken into account.

  However, we have to consider that the Test Bed 2 switching function was implemented on a quite limited scale, which might appear a bit too simplistic to evaluators used to working very large systems.

  The evaluators recommended to draw the cameras on the maps in a much more realistic way, which is a valuable comment.

  At one of the End-users companies, they use some keyboard shortcuts facilities and it seems to be very efficient.

  The four “sub-screens” on a VGA screen is considered as too limited in size and number and switching images on external monitors is recommended.

- ***On the efficiency of the search and retrieve functions***

  In general the score was quite good. Some very valuable suggestions have been made.

  Some comments of the evaluators deal with the understanding of some commands (today-now buttons, cancel-ok buttons, pause-stop buttons). It seems likely that their usage is less intuitive than for the other functions.

  If the bookmarking facility is to be implemented, the evaluators recommend a related filter be added to the query functions.

Another issue is related to the laws (duration of keeping the images stored before erasing and watermarking). Those comments are extremely relevant for a real implementation.

The additional facility of image export is considered as necessary in order to fully exploit the retrieval function.

- ***On the alarm management***

  A number of criticisms are about the way the alarms are presented. The comments made are all relevant.

  The requested reporting facility of the alarms generated is given in any system generating alarms. This should be the case in ADVISOR too.

  The request of the evaluators to be able to move quickly from live alarm images to the immediate past images is a good idea. However, it means a sort of fast changing of working mode (from live to retrieve and vice-versa), which might cause some confusion after a while.

3.6.2.4 Actions and ideas to improve the operating of the ADVISOR system

- ***To improve the human control***

  Little needs to be done according to this second intermediate evaluation.

- ***To improve the efficiency of the switching functions***

  We have to complete larger scale implementation of the switching facilities, i.e. more stations and more cameras so that the prototype appears more credible.

  We should represent the cameras on the maps so that the operator could easily understand their field of view.

  Keyboard shortcuts could be envisaged if time is available.

  To switch images on external monitors should be possible in the future, but seems difficult to realise since it requires different hardware and additional time to implement.

- ***To improve the efficiency of the search and retrieve functions***

  It is necessary to clarify the use of the buttons mentioned by the evaluators, i.e. :

  - today/now buttons in the query window
  - cancel/ok buttons in the query window
  - pause/stop buttons in the VCR-like commands

  To fully implement the bookmarking function, including the query based on human operator annotation.

  Regarding the specifications needed to comply with rules or laws, they should be implemented for a real deployment of the system. At the level of the ADVISOR prototype, these points should somehow be addressed, i.e. :

- watermarking
- automatic erasing of the recorded images after a programmed duration

The possibility to export images in an open format should be envisaged.

- ***To improve the alarm management***

  We should improve the presentation of the alarm messages, i.e. :

  - only one line per alarm, with colour changing according to the status of the alarm (active, not-active, acknowledged, not acknowledged)
  - to start each line with date and time information
  - to mask the information that is not useful for the operator
  - to increase the number of available simultaneous alarms visible in the window

  We should foresee some function to export the list of the alarms generated. Due to the limited time available, probably a simple file would be preferred at this stage.

  We should investigate which modification of the specification would be needed to go quickly from a live situation to a retrieve situation (and vice-versa) without being forced to use the search mode window. Great attention should be paid to ensure that the use of the system does not become confusing. At this stage of the project however, it is unlikely that this request could be implemented.

3.6.3 Using criteria

3.6.3.1 Table of individual evaluations of using

Using criteria	STIB 1	STIB 2	TMB 1	TMB 2
Clearness of information / messages Are all the messages that are provided by the system simple, clear and unambiguous ?	C	D	C	C
Adaptation to users' understanding Are all the messages communicated to and actions required from operators well adapted to their level of understanding ?	C	C	C	C
Easiness to learn Is it easy to learn and use the system ?	C	C	C	D
General ergonomics of the HCI Is the ADVISOR system user-friendly ? Is the ADVISOR system pleasant to use ?	B	C	D	D

Using criteria	STIB 1	STIB 2	TMB 1	TMB 2	Mean
Clearness of information / messages Are all the messages that are provided by the system simple, clear and unambiguous ?	0,65	0,95	0,65	0,65	0,725
Adaptation to users' understanding Are all the messages communicated to and actions required from operators well adapted to their level of understanding ?	0,65	0,65	0,65	0,65	0,65
Easiness to learn Is it easy to learn and use the system ?	0,65	0,65	0,65	0,95	0,725
General ergonomics of the HCI Is the ADVISOR system user-friendly ? Is the ADVISOR system pleasant to use ?	0,35	0,65	0,95	0,95	0,725

3.6.3.2 Additional comments on using of the system

Comments provided by the evaluators and collected from the summarising discussion held after the individual evaluations :

- ***Clearness of information / messages***
  - Are all the messages that are provided by the system simple, clear and unambiguous ?
    - ✍ The wording "retrieve" to name the information box is not well suited
    - ✍ Help is required to understand colour code for alarms and images on the screens

- ***Easiness to learn***
  - o Is it easy to learn and use the system ?
    - ✍ Very easy from general opinion

- ***General ergonomics of the HCI***
  - o Is the ADVISOR system user-friendly ?
  - o Is the ADVISOR system pleasant to use ?
    - ✍ The “close” button in selection and search modes is too small
    - ✍ The “cancel” and “ok” buttons in the search mode should be replaced by something more clear, for example ☒ or “quit” instead of “cancel” and “confirm” instead of “ok”
    - ✍ VCR-like commands in the retrieve mode are not very easy to use correctly. For example “eject” should be better than “stop”

3.6.3.3 Review of the evaluation results for the using criteria

- ***On the clearness of the information and messages***

  Overall, the evaluators gave good marks for this point. However, we have to take the comments made in previous criteria into account (see sections 3.6.1.2 and 3.6.2.2).

  Moreover, they provided some good suggestions to ease the understanding of the systems working.

- ***On the adaptation to the users’ understanding***

  In general they all agreed that the system was not very difficult to understand. However, taking the potential skill levels of real operators into account one has to be cautious.

- ***On the easiness to learn the system***

  Although the learning phase has been quite short, all the evaluators agreed that the system was very easy to learn.

- ***On the general ergonomics of the HCI***

  The evaluators who are used to work with similar systems had taken advantage of the Windows paradigm and found the HCI overall very user-friendly. This is also reflected in the evaluation of the easiness to learn it.

  Some additional suggestions have however been submitted in order to improve the ergonomic of the HCI.

3.6.3.4 Actions and ideas to improve the using of the ADVISOR system

- ***To improve the clearness of the information and messages***

  We should make the colour code used for the alarm messages much more clear to understand.

  We should modify some windows naming, in particular on the left part of the HCI screen.

  Obviously, we should implement the improvements mentioned at sections 3.6.1.4 and 3.6.2.4 above.

- ***To improve the users' understanding***

  Above the other recommendations made by the evaluators, we should ideally translate all the messages provided by the system into the mother language of the operators.

- ***To improve the easiness to learn the system***

  Nothing more to do than what is mentioned elsewhere which should also facilitate the learning of the system.

- ***To improve the general ergonomic of the HCI***

  Above the recommendations made before, we should :

  - enlarge the close button in selection and search mode
  - replace the "cancel" button in search mode by "undo" (instead of "quit" as originally suggested)
  - replace the "ok" button in search mode by "start" because it starts the retrieve of the selected sequences (instead of "confirm" as originally suggested)
  - replace the "stop" button in the VCR-like commands by "live" because it returns back to the live mode (instead of "eject" as originally suggested)

4 CONCLUSIONS

The four evaluators from both STIB and TMB were interested by the ADVISOR system although it was still an intermediate version.

Two of them had never seen the system before, and two of them participated in the first intermediate evaluation (after Test Bed 1). No significant difference appears between the evaluation of both groups.

For this reason we believe that their comments and suggestions are objective and not influenced by their possible previous understanding of the project.

This second intermediate evaluation was both quantitative and qualitative and has been very valuable to improve a number of issues, especially on what concerns the Human Computer Interface specifications. This will lead us to update Deliverable R5.1 on HCI Specifications.

As expected too, this second intermediate evaluation provided feed-back to the core ADVISOR software that will be taken into consideration in view of the final demonstration.