ADVISOR

R7.2 Technical Validation of Test-Bed 2

IST-1999-11287

Classification: Public

Author A. D. Barnett

Date 21 February 2003

The ADVISOR Consortium consists of:

Thales Research and Technology - UK	Co-ordinator	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

©Thales Research and Technology - UK 2001-2003

Thales Research and Technology - UK,
Worton Drive,
Worton Grange Industrial Estate,
Reading,
Berkshire
RG2 0SB

This document is issued by Thales Research and Technology - UK (hereinafter called TRT(UK)) in confidence and is not to be reproduced in whole or in part without the prior written permission of TRT(UK). The information contained herein is the property of TRT(UK) and is to be used for the purpose for which it is submitted and is not to be released in whole or in part or the contents disclosed to a third party without the prior written permission of TRT(UK).

Classification: Public

Documents that are classified as Internal may not be distributed outside the Project Team, without prior written consent of all members of the Steering Committee.

MODIFICATION RECORD

Version	Date	Author	Purpose/Description
Draft A	09-01-2003	A.D. Barnett	Initial draft of TB2++ validation report
Draft B	17-02-2003	A.D. Barnett	Updated draft of TB2++ validation report incorporating review comments
Issue 1	20-02-2003	A.D. Barnett	Issue 1 for Release

Contents

Tables

Executive Summary

This document describes the validation of the Annotated Digital Video for Intelligent Surveillance and Optimised Retrieval (ADVISOR) system Test-Bed 2 [1]. The report includes the data used for the validation, the validation process employed, the results obtained and an analysis of the results.

The ADVISOR system was tested in two-stages. The first stage of validation, involved playing a sequence of footage from a CCTV camera, through the system and noting the resulting behaviour report generated by ADVISOR. If the system generated the correct behaviour report then the second-stage of validation was applied which measured the accuracy of the report. This was achieved by measuring the percentage of the sequence during which the correct report was generated.

The system was tested using various sequences involving the behaviours of blocking and fighting, as described in [1]. Blocking corresponds to a situation when a group of at least 2 people is stopped in a predefined zone for at least 4 seconds and can potentially block the path of other people. Fighting corresponds to a situation when a group of people (at least 2 persons) are pushing, kicking or grasping each other for at least 2 seconds.

Single camera sequences were mainly used but a combined sequence using two cameras in parallel was also tested.

It was found that out of a total of 55 single camera sequences, there were 22 (40%) true results, including the true negative results, where no events were generated for sequences that had no fighting or blocking. The blocking sequences were found to be correctly reported more often than the fighting sequences. Indeed, in all four individual blocking sequences, blocking was correctly reported and in the consecutive sequence, blocking was also correctly reported each time, until the system crashed due to an overflow failure.

Fighting was reported less successfully, with 80 % of the individual sequences being correctly reported but only 15% of the consecutive sequences being correctly identified. This is due to a problem with background files, where the system is unable to reset itself to an empty scene after fighting has occurred and will be rectified in the demonstrator system.

For the sequences that were correctly reported in the first-phase of validation, the behaviour reports were found to be very accurate under the criteria of the second-phase of validation. Overall, those sequences tested in the second-phase gave an average accuracy score of 82%.

When testing ADVISOR with two cameras in combination, no behaviour reports were generated, even though there were spurious overcrowding reports generated when the sequences were run through individually. This may indicate a problem with multi-camera use.

In conclusion, the ADVISOR system meets the requirements of Test-Bed 2 as laid out in the functional specification document [1]. The system works very well on a few individual sequences, but does not work so well on other sequences, even when the fighting is quite exaggerated. In particular, the fighting algorithms have low success rate in recognising fighting behaviour. On the other hand, the blocking algorithms seem to work very well.

1 Introduction

This document describes the validation of the Annotated Digital Video for Intelligent Surveillance and Optimised Retrieval (ADVISOR) system under Test-Bed 2 [1]. The report includes the data used for the validation, the validation process employed, the results obtained and an analysis of the results.

The requirements of Test-Bed 2 as described in [1], validated in this report are as follows;

- Test-Bed 2 startup, including implementation of configuration files
- Test-Bed 2 close down
- Source material requirements
- Capture module functionality
- Machine vision processing functionality, comprising recognition of blocking and fighting
- Archive, Search and Retrieval functionality
- Human Computer Interface (HCI) functionality

Although the requirements of Test-Bed 2 are not all tested individually, their functionality is observed during the validation of the whole ADVISOR system. In particular, the overall operation of the Capture, the Archive, Search and Retrieval (ASR) and the Human Computer Interface (HCI) modules is tested during the validation process.

The ADVISOR system detects the motion of people by CCTV cameras and attempts to classify and report their behaviour. Identification of these behaviours in the system generates an alarm at the user interface.

The behaviours tested in Test-Bed 2 are those of blocking and fighting, as described in [1]. Blocking corresponds to a situation when a group of at least 2 people is stopped in a predefined zone for at least 4 seconds and can potentially block the path of other people. Fighting corresponds to a situation when a group of people (at least 2 persons) are pushing, kicking or grasping each other for at least 2 seconds.

Initially, the system is intended for use on underground Metro systems and is tested using sequences of CCTV footage from Yzer station on the STIB (Brussels Metro) and from the Sagrada Familia station on the TMB (Barcelona Metro).

The sequences of CCTV footage are firstly ground-truthed by identifying the behaviour by inspection and recording the times at which those behaviours occurred. The results obtained when the sequence is played through the system are then compared with the ground-truth.

The ADVISOR system is tested here in two-stages. The first stage of validation, involves playing a sequence of footage from a CCTV camera, through the system and noting the resulting behaviour report generated by ADVISOR. If the system generates the correct behaviour report then the second-stage of validation can be applied which measures the accuracy of the report. This is achieved by measuring the percentage of the sequence during which the correct report is generated.

The system is tested using various sequences involving the behaviours of blocking and fighting. Single camera sequences are mainly used but a combined sequence using two cameras in parallel is also tested.

The main limitations of the validation process are the time required to test the system and the amount of test data available. Whilst more than fifty sequences have been validated here, a larger number of sequences would have given a more accurate validation of the system. However, the system needs to be reset between each individual run and between sets of runs, and requires the

reconfiguration of the cameras in the Capture and Human Computer Interface (HCI) modules, the reconfiguring of the background image for each sequence and the clearance of the archive. The system then needs to be started up and the sequence played through the system. This is a time consuming process. The test data used in the validation was recorded specifically for ADVISOR, mainly using actors to enact the required behaviours. Most of these recorded sequences have been used in the validation process. However, more sequences, containing different types of fighting, for example, would give a more accurate appraisal of the performance of ADVISOR. The recording and digitisation of more acted sequences is difficult to co-ordinate and is time consuming.

1.1 Document Layout

In the following section, the validation process is described, followed by a description of the data used and the ground-truth of that data in Section 3. In Section 4, the results of the response of the ADVISOR system to the data is presented and analysed. Finally, in Section 5 a summary of the report and the conclusions drawn are given.

2 Validation Process

2.1 Validation Process Details

The validation process is carried out in two stages. Firstly, a simple validation process is used to test the system, whereby a true or false result is obtained if the system has detected the observed event or not. This includes any false reporting of events, i.e. the system reporting 'fighting' when no fighting has occurred.

The second stage of validation is applied to sequences where the system has generated a true result in the first stage. The time at which the event is reported, and the duration of the event, is compared with the ground-truth, giving, as the result of the second-stage validation, a percentage of the sequence during which the correct event was reported.

2.1.1 First-Stage Validation Process

For the first-stage validation process, only short sequences are used. For simple validation, an alarm produced anywhere during the sequence is recorded simply as the fact that the alarm was generated. The ground-truth data is produced in a similar way - see Section 3.1.1. Simple validation consists of a binary comparison for each alarm (and event, if events are validated) between the actual behaviour and the ground-truth data.

An example is probably necessary. Consider a sequence SEQ-NNN for C02. The ground-truth data for this sequence might be:

Fighting: yes
Blocking: no

In other words, the sequence *should* generate a 'fighting' alarm and should *not* generate a 'blocking' alarm. The sequence is run through the test-bed, and the actual performance is recorded to the same level of detail. That is, the HCI is used to report any alarms generated and the results are recorded. For example, they might be:

Fighting: no
Blocking: no

In other words, the sequence *did not* generate a 'fighting' alarm and *did not* generate a 'blocking' alarm. The ground-truth and actual are then compared, which in this example would give:

Fighting: false negative (fn)
Blocking: true negative (tn)

which means that 'fighting' was not detected, but should have been - that is, it was incorrectly or falsely detected - and 'blocking' was not detected, and should not have been - that is, it was correctly not detected. There are, in general, four possible outcomes:

tp - true positive: the alarm *should* have been generated, and it *was*.
tn - true negative: the alarm should *not* have been generated, and it was *not*.
fp - false positive: the alarm should *not* have been generated, but it *was*.
fn - false negative: the alarm *should* have been generated, but it was *not*

The basic behaviour and ground-truth is assessed for single-camera and two-camera sequences only.

2.1.2 Second-Stage Validation Process

The second-stage of validation involves measuring the accuracy of a true positive result obtained at the first stage. Quite simply, the percentage of a sequence during which the ADVISOR system produces correct behaviour reports is measured. This includes true negative responses. So, for example, if ADVISOR reports a sequence as having fighting for 45 seconds, when the ground-truth shows that 60 seconds of fighting occurred, then a score of 75% would be awarded. The score also includes true negative periods of the sequence, i.e. if nothing happens and no reports are generated then the report is counted as being correct. A delay of 5 seconds after the behaviour starts in the ground-truth is permitted in the measurement as this is the specified delay in the behaviour detection algorithms.

3 Ground-Truth Data

3.1.1 First-Stage Validation Ground-Truth Data

The first-stage ground-truth is assessed for single-camera and two-camera sequences only.

An example is probably necessary. Consider a sequence SEQ-NNN for C02. The ground-truth data for this sequence might be:

Fighting: yes
Blocking: no

In other words, the sequence *should* generate a 'fighting' alarm and should *not* generate a 'blocking' alarm.

The first-stage ground-truth data is created by *inspection*. That is, a competent authority examines the sequence and decides on what behaviours are being exhibited. This has two main weaknesses:

1) Definition - what behaviour should cause the alarm?
2) Interpretation - is the behaviour exhibited?

Even given a suitable definition of what constitutes the behaviour that should produce an alarm, it is still a subjective assessment whether or not a particular sequence does in fact meet the criteria.

The view before validation was that this should not be a problem, because alarms should be high-level ideas with human-understood meanings. As such, the user should easily understand them. This should actually help with producing useful ground-truth data.

For example, consider 'fighting'. The user (an operator) is not going to be concerned with the technical definition of 'fighting'. They are going to judge the system's performance quite simply: how often does a 'fighting' alarm actually show what *they* consider to be fighting, and how often is what *they* would call fighting actually ignored?

The definition of 'blocking', has been clarified in [1], and has been used in the ground-truthing of the sequences.

The sequences used in the first-stage of the validation process are given in Table 1. The following sequences were run individually with the system being reset between each run:

- SEQ-017-C05: STIB (blocking)
- SEQ-003-C05: STIB (fighting)
- SEQ-003-C10: STIB (fighting)
- SEQ-062-C10: TMB (fighting)
- SEQ-054-C12: TMB (blocking)
- SEQ-046-C10: TMB (blocking)
- SEQ-009-C05: STIB (fighting)
- SEQ-035-C07: STIB (blocking)

The following sets of sequences were run consecutively, without resetting between individual sequences (the system was reset between each set of sequences):

- SEQ-046-C12 to SEQ-058-C12: TMB (blocking)
- SEQ-060-C10 to SEQ-066-C10: TMB (fighting)

- SEQ-059-C12 to SEQ-066-C12: TMB (fighting)
- SEQ-073-C01 to SEQ-086-C01: TMB (fighting)
- SEQ-108-C02 to SEQ-113-C02: STIB (fighting)
- SEQ-108-C04 to SEQ-113-C04: STIB (fighting)
- SEQ-108-C02 to SEQ-113-C02 and SEQ-108-C04 to SEQ-113-C04 combined: STIB (fighting)

The background image used for each sequence is given in [2]. In the table, a tick (✓) indicates that the behaviour occurred during the sequence, while a cross (×) indicates that the behaviour did not occur in the sequence.

SEQ	Ext Ref	Behaviour	Station and Camera							
			SGFM_C01	SGFM_C10	SGFM_C12	YZER_C02	YZER_C04	YZER_C05	YZER_C07	YZER_C10
SEQ-017	1.1.02	blocking						✓		
SEQ-035	1.4.02	blocking							✓	
SEQ-046	SGFM 9.1.01	blocking		✓	✓					
SEQ-047	SGFM 9.1.02	blocking			✓					
SEQ-048	SGFM 9.1.03	blocking			✓					
SEQ-049	SGFM 9.1.04	blocking			✓					
SEQ-050	SGFM 9.1.05	blocking			✓					
SEQ-051	SGFM 9.1.06	blocking			✓					
SEQ-052	SGFM 9.1.07	blocking			✓					
SEQ-053	SGFM 9.1.08	blocking			✓					
SEQ-054	SGFM 9.1.09	blocking			✓					
SEQ-055	SGFM 9.1.10	blocking			✓					
SEQ-056	SGFM 9.1.11	blocking			✓					
SEQ-057	SGFM 9.1.12	blocking			✓					
SEQ-058	SGFM 9.1.13	blocking			✓					
SEQ-003	2.1.01	fighting						✓		✓
SEQ-009	2.1.05	fighting						✓		✓
SEQ-059	SGFM 10.1.01	fighting			✓					
SEQ-060	SGFM 10.1.02	fighting		✓	✓					
SEQ-061	SGFM 10.1.03	fighting		✓	✓					
SEQ-062	SGFM 10.1.04	fighting		✓	✓					
SEQ-063	SGFM 10.1.05	fighting		✓	✓					
SEQ-064	SGFM 10.1.06	fighting		✓	✓					
SEQ-065	SGFM 10.1.07	fighting		✓	✓					
SEQ-066	SGFM 10.1.08	fighting		✓	✓					
SEQ-078	SGFM 10.2.07	fighting	✓							
SEQ-080	SGFM 10.2.09	fighting	✓							
SEQ-081	SGFM 10.2.10	fighting	✓							
SEQ-082	SGFM 10.2.11	fighting	✓							
SEQ-083	SGFM 10.2.12	fighting	✓							
SEQ-084	SGFM 10.2.13	fighting	✓							
SEQ-085	SGFM 10.2.14	fighting	✓							
SEQ-086	SGFM 10.2.15	fighting	✓							
SEQ-108	2.2.12	fighting				✓	✓			
SEQ-109	2.2.13	fighting				✓	✓			
SEQ-110	2.2.14	fighting				×	×			
SEQ-111	2.2.15	fighting				×	×			
SEQ-112	2.2.16	fighting				×	×			
SEQ-113	2.2.17	fighting				×	×			

Table 1 – First-Stage Ground-Truth Data – Blocking and Fighting

3.1.2 Second-Stage Validation Ground-Truth Data

The second-stage ground-truth validation data for the sequences which produce a true positive result, is not given here but instead is presented in the following section with the results of the second-stage validation. It is unnecessary to ground-truth all the sequences at the second-level of

validation because not all of the sequences are tested in the second stage.

4 Validation Results and Analysis

4.1.1 First-Stage Validation Results

The response of the ADVISOR system to the sequences given in Table 1 is given in Table 2 below. True positive results are indicated with a tick and false negative results with a cross. False positive events are also indicated in the table by the abbreviation *fp* and true negative results by *tn*. The abbreviation *tp+fp* indicates that the correct behaviour was reported during the sequences, but an incorrect behaviour was also reported during the sequence. Similarly, the abbreviation *fn+fp* indicates that the correct behaviour was not reported during the sequence and an incorrect behaviour was also reported during the sequence.

The highlighted columns indicate where sequences have been run consecutively without resetting the system, rather than individually.

SEQ	Ext Ref	Behaviour	Station and Camera							
			SGFM _C01	SGFM _C10	SGFM _C12	YZER _C02	YZER _C04	YZER _C05	YZER _C07	YZER _C10
SEQ-017	1.1.02	blocking	-	-	-	-	-	tp	-	-
SEQ-035	1.4.02	blocking	-	-	-	-	-	-	tp +fp	-
SEQ-046	SGFM 9.1.01	blocking	-	tp	-	-	-	-	-	-
SEQ-054	SGFM 9.1.09	blocking	-	-	tp +fp	-	-	-	-	-
SEQ-046	SGFM 9.1.01	blocking	-	-	tp	-	-	-	-	-
SEQ-047	SGFM 9.1.02	blocking	-	-	tp	-	-	-	-	-
SEQ-048	SGFM 9.1.03	blocking	-	-	tp +fp	-	-	-	-	-
SEQ-049	SGFM 9.1.04	blocking	-	-	tp	-	-	-	-	-
SEQ-050	SGFM 9.1.05	blocking	-	-	tp	-	-	-	-	-
SEQ-051	SGFM 9.1.06	blocking	-	-	-	-	-	-	-	-
SEQ-052	SGFM 9.1.07	blocking	-	-	-	-	-	-	-	-
SEQ-053	SGFM 9.1.08	blocking	-	-	-	-	-	-	-	-
SEQ-054	SGFM 9.1.09	blocking	-	-	-	-	-	-	-	-
SEQ-055	SGFM 9.1.10	blocking	-	-	-	-	-	-	-	-
SEQ-056	SGFM 9.1.11	blocking	-	-	-	-	-	-	-	-
SEQ-057	SGFM 9.1.12	blocking	-	-	-	-	-	-	-	-
SEQ-058	SGFM 9.1.13	blocking	-	-	-	-	-	-	-	-
SEQ-003	2.1.01	fighting	-	-	-	-	-	tp	-	tp
SEQ-009	2.1.05	fighting	-	-	-	-	-	tp	-	-
SEQ-062	SGFM 10.1.04	fighting	-	tp	fn	-	-	-	-	-
SEQ-059	SGFM 10.1.01	fighting	-	-	fn	-	-	-	-	-
SEQ-060	SGFM 10.1.02	fighting	-	fn	fn +fp	-	-	-	-	-
SEQ-061	SGFM 10.1.03	fighting	-	fn	tp	-	-	-	-	-
SEQ-062	SGFM 10.1.04	fighting	-	tp	fn	-	-	-	-	-
SEQ-063	SGFM 10.1.05	fighting	-	fn	tp +fp	-	-	-	-	-
SEQ-064	SGFM 10.1.06	fighting	-	fn +fp	fn +fp	-	-	-	-	-
SEQ-065	SGFM 10.1.07	fighting	-	tp	fn +fp	-	-	-	-	-
SEQ-066	SGFM 10.1.08	fighting	-	fn +fp	fn	-	-	-	-	-
SEQ-073	SGFM 10.2.02	fighting	fn	-	-	-	-	-	-	-
SEQ-074	SGFM 10.2.03	fighting	fn	-	-	-	-	-	-	-
SEQ-075	SGFM 10.2.04	fighting	fn	-	-	-	-	-	-	-
SEQ-076	SGFM 10.2.05	fighting	fn	-	-	-	-	-	-	-
SEQ-077	SGFM 10.2.06	fighting	fn	-	-	-	-	-	-	-
SEQ-078	SGFM 10.2.07	fighting	fn	-	-	-	-	-	-	-
SEQ-079	SGFM 10.2.08	fighting	fn	-	-	-	-	-	-	-
SEQ-080	SGFM 10.2.09	fighting	tp	-	-	-	-	-	-	-
SEQ-081	SGFM 10.2.10	fighting	fn	-	-	-	-	-	-	-
SEQ-082	SGFM 10.2.11	fighting	fn	-	-	-	-	-	-	-
SEQ-083	SGFM 10.2.12	fighting	fn	-	-	-	-	-	-	-
SEQ-084	SGFM 10.2.13	fighting	fn	-	-	-	-	-	-	-
SEQ-085	SGFM 10.2.14	fighting	fn	-	-	-	-	-	-	-
SEQ-086	SGFM 10.2.15	fighting	fn	-	-	-	-	-	-	-

SEQ	Ext Ref	Behaviour	Station and Camera							
			SGFM _C01	SGFM _C10	SGFM _C12	YZER _C02	YZER _C04	YZER _C05	YZER _C07	YZER _C10
SEQ-108	2.2.12	fighting	-	-	-	fn	fn	-	-	-
SEQ-109	2.2.13	fighting	-	-	-	fn	fn	-	-	-
SEQ-110	2.2.14	fighting	-	-	-	tn	tn	-	-	-
SEQ-111	2.2.15	fighting	-	-	-	fp	fp	-	-	-
SEQ-112	2.2.16	fighting	-	-	-	tn	fp	-	-	-
SEQ-113	2.2.17	fighting	-	-	-	tn	fp	-	-	-

Table 2 – First-Stage Actual Data – Blocking and Fighting, Single Camera Sequences

SEQ	Ext Ref	Behaviour	Station and Camera							
			SGFM _C01	SGFM _C10	SGFM _C12	YZER _C02	YZER _C04	YZER _C05	YZER _C07	YZER _C10
SEQ-108	2.2.12	fighting	-	-	-	✕	✕	-	-	-
SEQ-109	2.2.13	fighting	-	-	-	✕	✕	-	-	-
SEQ-110	2.2.14	fighting	-	-	-	tn	tn	-	-	-
SEQ-111	2.2.15	fighting	-	-	-	tn	tn	-	-	-
SEQ-112	2.2.16	fighting	-	-	-	tn	tn	-	-	-
SEQ-113	2.2.17	fighting	-	-	-	tn	tn	-	-	-

Table 3 – First-Stage Actual Data – Blocking and Fighting, Two-Camera Sequences

4.1.2 First-Stage Validation Analysis

It can be seen from Table 2 that out of a total of 55 single camera sequences, there were 22 (40%) true results, including the true negative results.

Overall, out of the 9 individual sequences, comprising both blocking and fighting, 8 correct reports were generated, giving a success rate of 89%. Of the remaining 46 sequences in four different sets, correct reports were generated 14 times, giving a success rate of 30%.

In the four individual blocking sequences tested, the ADVISOR system correctly reported blocking in all four cases, with one sequence giving an additional false positive report of fighting. In the functional specification for Test-Bed 2 it states that "The locations that have been configured as capable of triggering a 'blocking' alarm in Test Bed 2 are … Yzer camera C05 and Yzer camera C07." The results show that blocking was reported in cameras C10 and C12 in SGFM, which are correct reports but they exceed the functional specification of Test Bed 2. It has been pointed out by M. Renard (Vigitec), that this is a mistake in the functional specification for Test Bed 2 because blocking events in the field of view of C10 and C12 were recorded for this purpose.

Whilst running the blocking sequences SEQ-046-C12 to SEQ-058-C12 the system reported blocking correctly (but again exceeded the functional specification) until the system crashed during SEQ-050-C12 due to an *overflow* failure.

In the five individual fighting sequences tested, the ADVISOR system correctly reported fighting in four (80%) of the cases. These sequences are analysed further in the second-stage of validation. Out of a total of 38 fighting sequences, ADVISOR correctly produced nine fighting reports, giving a success rate of 24%. Out of the 33 fighting sequences in 4 different sets, ADVISOR only produced the correct report 5 times, giving a success rate of 15%.

In sequences SEQ-060-C10 to SEQ-066-C10 and SEQ-059-C12 to SEQ-066-C12, blocking was erroneously reported during the fighting scenes. Once again this violates the functional specification of Test Bed 2 as well as being an incorrect behaviour report.

In sequences SEQ-108-C02 to SEQ-113-C02 and SEQ-108-C04 to SEQ-113-C04, run

consecutively and in combination, no reports were produced at all, despite the fighting scenes in the first two sequences. However, the erroneous reports of overcrowding reported by camera C04 in the single camera test were not reported in the two-camera test, giving true negative results for the remaining sequences.

In the validation of the two-camera sequences, no behaviour reports, either events or alarms were generated, whereas in the single camera runs of the same sequences, erroneous overcrowding and fighting reports were generated. This requires further investigation.

4.1.3 Second Stage Validation Results

The second stage validation results for each sequence that produced a true positive result at the first stage of validation are presented as follows.

Sequence	Behaviour	Sequence duration (s)	Ground-Truth	Actual response	% Accuracy
SEQ-017	Blocking	119	Blocking: 22:24:46 to 22:26:27	Blocking: 22:24:48 – end	100%
SEQ-003-C05	Fighting	63	Fighting: 00:18:12 to 00:18:28	Fighting: 00:18:17 to 00:18:29	100%
SEQ-003-C10	Fighting	63	Fighting: 00:17:50 to 00:17:52	Fighting: 00:17:55 to 00:17:57	100%
SEQ-062-C10	Fighting	33	Fighting: 00:45:49 to 00:46:00	Fighting: 00:45:54 to 00:46:03	100%
SEQ-054-C12	Blocking	70	Blocking: 22:55:05 to 22:55:40	Blocking: 22:55:13 to 22:55:22 Blocking: 22:55:13 to 22:55:43	93%
SEQ-046-C10	Blocking	98	Blocking: 22:36:39 to 22:37:49	Blocking: 22:36:48 to 22:37:22 Blocking: 22:36:48 to 22:37:23 Blocking: 22:36:48 to 22:37:39 Fighting: 22:37:39 to 22:37:45 Blocking: 22:36:48 to end Fighting: 22:37:39 to 22:37:51 Blocking: 22:36:48 to end	93% (ignoring the erroneous fighting reports)
SEQ-009-C05	Fighting	59	Fighting: 00:23:55 to 00:24:07	Fighting: 00:24:07 to 00:24:14	85%
SEQ-035-C07	Blocking	142	Blocking: 22:57:21 to 22:59:10	Blocking 22:57:20 to 22:57:48 Blocking 22:57:20 to 22:59:09 Blocking 22:57:20 to end Fighting 22:59:13 to end	96% (ignoring the erroneous fighting report)

Sequence	Behaviour	Sequence duration (s)	Ground-Truth	Actual response	% Accuracy
SEQ-046-C12	Blocking	98	Blocking: 22:36:39 to 22:37:49	Blocking: 22:36:52 to 22:36:55 Blocking: 22:36:52 to 22:36:58 Blocking: 22:36:52 to 22:37:31	73%
SEQ-047-C12	Blocking	113	Blocking: 22:38:22 to 22:39:44	Blocking: 22:38:50 to 22:38:53 Blocking: 22:38:50 to 22:39:19 Blocking: 22:38:50 to end	73%
SEQ-048-C12	Blocking	85	Blocking: 22:40:10 to 22:41:03	Fighting: 22:39:49 to end Blocking: 22:38:50 to 22:41:11	75% (ignoring the erroneous fighting reports)
SEQ-049-C12	Blocking	105	Blocking: 22:41:42 to 22:42:53	Blocking: 22:38:50 to 22:41:23 Blocking: 22:41:42 to 22:42:54	100%
SEQ-050-C12 (run consecutively	Blocking	59	Blocking: 22:43:26 to 22:44:00	Blocking: 22:41:42 to 22:43:03	66%
SEQ-061-C12 (run in a consecutive sequence)	Fighting	33	Fighting: 00:43:33 to 00:44:40	Fighting 00:43:33 to 00:43:43	100%
SEQ-063-C12	Fighting	42	Fighting: 00:44:43 to 00:44:54	Blocking 00:44:49 to 00:44:49 Blocking 00:44:49 to 00:44:56 Fighting 00:44:58 to 00:45:02	69% (ignoring the erroneous blocking reports)
SEQ-080-C01	Fighting	110	Fighting: 00:10:52 to 00:11:20	Fighting 01:11:42 to 01:11:50	0%

Table 4 – First-Stage Actual Data – Blocking and Fighting, Single Camera Sequences

4.1.4 Second Stage Validation Analysis

It should first be noted that in the calculation of the percentage accuracy, a delay of up to five seconds after the behaviour began or finished in the ground-truth, was permitted when scoring the accuracy of the actual response. For example, if an event was ground-truthed to start at 10 seconds, and the ADVISOR system reported the event as starting at 15 seconds then that was scored to be correct.

Overall, from the results tabulated in Table 4, it can be seen that where ADVISOR generated correct behaviour reports and alarms in the first stage of validation, the accuracy of the reports as measured in the second-stage of validation were quite high. Indeed, 10 out of the 16 sequences produced behaviour reports that were scored to be above 80% accurate, compared with the ground-truth. With the exception of one sequence, the remaining sequences all produced reports

that were greater than 66% accurate. This shows that when ADVISOR recognises behaviour correctly then it is quite accurate in the timing of the reports.

Overall, the sequences tested in the second stage of validation produced an average accuracy of 83%, which can be considered to be quite high.

In sequence SEQ-046-C10, which shows blocking at SGFM, there is a spurious report of fighting. In fact, by inspection of the sequence it can be seen that the actors in the sequence are greeting and kissing each other as they are leaving the scene. It is therefore understandable that this close contact may be incorrectly identified as fighting behaviour by the behaviour recognition algorithms.

From the second-stage validation results there seems to be no difference in the accuracy of the blocking or fighting events, i.e. it cannot be said that fighting is more accurately identified than blocking or vice versa.

5 Summary and Conclusions

In this report, the ADVISOR system has been validated under Test-Bed 2 [1]. Only the behaviours of fighting and blocking have been investigated, as described in the functional specification and only single-camera and two camera sequences have been used.

In response to the functional specification of Test-Bed 2, given in [1] the following requirements were met under the validation. Whilst the requirements were not all tested individually, their functionality was observed during the validation of the whole ADVISOR system

- Test Bed 2 startup, including implementation of configuration files
- Test-Bed 2 close down
- Source material requirements
- Capture module functionality
- Machine vision processing functionality, comprising recognition of blocking and fighting
- Archive, Search and Retrieval functionality
- Human Computer Interface (HCI) functionality

The first comment to be made is on the robustness of the ADVISOR system as a whole. An attempt to perform the validation of ADVISOR under Test-Bed 2 was attempted previously. However, problems with the stability of the system meant that few sequences were able to be tested. When sequences were able to be run through the system it was found that a lot of false positive reports were generated. In this re-validation of Test-Bed 2, the ADVISOR system was found to be very robust with very few crashes. Indeed, the only major system failure occurred when an overflow fault was reported by the behaviour recognition module during a blocking sequence for the Sagrada Familia station. Even in this case, the archive module stopped but did not crash. The majority of individual sequences lasted for 3 to 5 minutes, while the consecutively run sequences lasted for approximately 15 minutes.

In the validation, a number of CCTV footage sequences were run through the ADVISOR system and the resulting behaviour reports and alarms recorded. The generated event reports were compared with the ground-truth of the scenarios by human inspection. Two stages of validation were applied. Firstly, a simple pass or fail criteria was applied where a pass was given if the ADVISOR system correctly reported the behaviour as identified in the ground-truth. If a sequence passed the first stage, a second stage of validation was applied, whereby the accuracy of the behaviour report was measured. This was achieved by measuring the percentage of the sequence in which ADVISOR correctly reported the events in the sequence, including periods where no events occurred. It should be stressed that the validation results are subjective, i.e. someone else performing the validation might obtain slightly different results, but overall, the start and finish of the behaviours displayed in the sequences are fairly clear.

It was found that out of a total of 55 single camera sequences, there were 22 (40%) true results, including the true negative results, where no events were generated for sequences that had no fighting or blocking. The blocking sequences were found to be correctly reported more often than the fighting sequences. Indeed, in all four individual blocking sequences, blocking was correctly reported and the consecutive sequence, blocking was also correctly reported each time, until the system crashed due to an overflow failure. Fighting was reported less successfully, with 80 % of the individual sequences being correctly reported but only 15% of the consecutive sequences being correctly identified. This is understood to be due to a problem with background files, where the system is unable to reset itself to an empty scene after fighting has occurred. This has been rectified in the demonstrator system.

For the sequences that were correctly reported in the first-phase of validation, the behaviour reports were found to be very accurate under the criteria of the second-phase of validation. Overall, those sequences tested in the second-phase gave an average accuracy score of 82%.

When testing ADVISOR with two cameras in combination, no behaviour reports were generated, even though there were spurious overcrowding reports generated when the sequences were run through individually. This may indicate a problem with multi-camera use.

In conclusion, the ADVISOR system meets the requirements of Test-Bed 2 as laid out in the functional specification document. The system works very well on a few individual sequences, but does not work so well on other sequences, even when the fighting is quite exaggerated. In particular, the fighting algorithms have low success rate in recognising fighting behaviour. On the other hand, the blocking algorithms seem to work very well.

5.1 Other comments on the operation of ADVISOR

Whilst validating the ADVISOR system, a few flaws and suggested improvements came to light. Firstly, when searching through the archive for events that had occurred in the TMB, there was no list of TMB stations in the HCI. This forced the selection of “All stations” when searching for events rather than the specific choice of the Sagrada Familia station. In addition, the list of available cameras is misleading and unhelpful. It would be of more use if the camera number were available for selection, such as C12 for example.

In the motion detector unit, when the system starts up a “EOF/read error” message is sometimes generated. It is believed that the error is generated by the choice of background image files. The details of this problem have not been investigated, i.e., it is not known which images cause the error messages to be generated, but the system does not crash, indeed it seems to operate normally.

Finally, the startup procedure is quite lengthy and fiddly. On average, it takes 10 to 15 minutes to reset the system, including clearing the archive and reconfiguring the cameras. Each individual module has to be registered manually with the CORBA naming service. It would be useful if there was a simple ‘on’ procedure, which automatically starts up the whole system.

6 References

[1] [AFS] ADVISOR-DOC-006, ADVISOR Functional Specification.

[2] Background Images Table,
ADVISOR_FTP/Project_Library/22_INRIA_Software_Releases/008_Background_Images/backgroundimagestable.htm

7 Glossary, Terms and Abbreviations

ADVISOR	Annotated Digital Video for Intelligent Surveillance and Optimised Retrieval
CCTV	Closed-Circuit Television
HCI	Human Computer Interface
SGFM	Sagrada Familia Metro station on the TMB
STIB	Société des Transports Intercommunaux de Bruxelles (Brussels Metro)
TMB	Transports Metropolitans de Barcelona (Barcelona Metro)
TRT(UK)	Thales Research & Technology (UK)
Yzer	Metro station on the STIB