



DPC: 19/30363050 DC

BSI Group Headquarters

389 Chiswick High Road London W4 4AL

Tel: +44 (0)20 8996 9000

Fax: +44 (0)20 8996 7400

www.bsigroup.com

Date: 16 July 2019

Origin: International

Latest date for receipt of comments: 04 September 2019

Project No. 2017/02145

Responsible committee: GW/1 Electronic security systems and products

Interested committees:

Title: Draft BS EN IEC 62676-6 Video surveillance systems for use in security applications. Part 6: Video content analytics – Performance testing and grading.

Please notify the secretary if you are aware of any keywords that might assist in classifying or identifying the standard or if the content of this standard

- i) has any issues related to 3rd party IPR, patent or copyright
- ii) affects other national standard(s)
- iii) requires additional national guidance or information

WARNING: THIS IS A DRAFT AND MUST NOT BE REGARDED OR USED AS A BRITISH STANDARD. THIS DRAFT IS NOT CURRENT BEYOND 04 September 2019

This draft is issued to allow comments from interested parties; all comments will be given consideration prior to publication. No acknowledgement will normally be sent. **See overleaf for information on the submission of comments.**

No copying is allowed, in any form, without prior written permission from BSI except as permitted under the Copyright, Designs and Patent Act 1988 or for circulation within a nominating organization for briefing purposes. Electronic circulation is limited to dissemination by e-mail within such an organization by committee members.

Further copies of this draft may be purchased from BSI Shop <http://shop.bsigroup.com> or from BSI Customer Services, Tel: +44(0) 20 8996 9001 or email cservices@bsigroup.com. British, International and foreign standards are also available from BSI Customer Services.

Information on the co-operating organizations represented on the committees referenced above may be obtained from <https://standardsdevelopment.bsigroup.com/>.

Responsible Committee Secretary: Miss Beverley Webb (BSI)
Direct tel: 0208 996 7061
E-mail: beverley.webb@bsigroup.com

Full document details: 19/30363050 DC

Introduction

This draft standard is based on national and international discussions. Your comments on this draft are invited and will assist in the preparation of the consequent standard.

For international standards, comments will be reviewed by the relevant UK national committee before sending the consensus UK vote and comments to the international committee, which will then decide appropriate action. If the international standard is approved, it is usual for the text to be published as a British Standard.

For national standards, comments will be reviewed by the relevant UK national committee and the resulting standards published as a British Standard.

UK Vote

Please indicate whether you consider the UK should submit a negative (with supporting technical reasons) or positive vote on this draft. Please indicate if you are aware of any reason why this draft standard should not be published as a British Standard.

Submission of Comments

- **Annotated drafts are not acceptable and will be rejected.**
- All comments should be submitted online at <https://standardsdevelopment.bsigroup.com/>. You will need to register in order to comment.

Template for comments and secretariat observations

Date: xx/xx/20xx	Document: ISO/DIS xxxx
------------------	------------------------

1	2	(3)	4	5	(6)	7
MB	Clause No./ Subclause No./Annex (e.g. 3.1)	Paragraph/Figure/ Table/Note	Type of comment	Commend (justification for change) by the MB	Proposed change by the MB	Secretariat observations on each comment submitted
	3.1	Definition 1	ed	Definition is ambiguous and needs clarifying.	Amend to read '...so that the mains connector to which no connection...'	
	6.4	Paragraph 2	te	The use of the UV photometer as an alternative cannot be supported as serious problems have been encountered in its use in the UK.	Delete reference to UV photometer.	

PROJECT NUMBER: IEC 62676-6 ED1	
DATE OF CIRCULATION: 2019-07-12	CLOSING DATE FOR COMMENTS: 2019-10-04
SUPERSEDES DOCUMENTS: 79/571/NP, 79/587/RVN	

IEC TC 79 : ALARM AND ELECTRONIC SECURITY SYSTEMS	
SECRETARIAT: France	SECRETARY: Mr Olivier Lanen
OF INTEREST TO THE FOLLOWING COMMITTEES:	PROPOSED HORIZONTAL STANDARD: ☐ Other TC/SCs are requested to indicate their interest, if any, in this CD to the secretary.
FUNCTIONS CONCERNED: ☐ EMC ☐ ENVIRONMENT ☐ QUALITY ASSURANCE ☐ SAFETY	

This document is still under study and subject to change. It should not be used for reference purposes.

Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

TITLE:
Video surveillance systems for use in security applications – Part 6: Video content analytics – Performance testing and grading

NOTE FROM TC/SC OFFICERS:

The TC 79 Officers wish to inform TC 79 that the project leader for 62676-6 has decided to step down from his position due to new priorities in his company.

In response to some discussions in WG 12 Mr. Ollencio D'Souza (AU) has volunteered to take over.

This proposal was accepted by the WG and by the TC 79 Officers. It is hereby submitted to TC 79 for endorsement. Any NCs disagreeing with this proposal and having another candidate are invited to make a counterproposal when commenting on the present CD.

NOTE FROM IEC CO:

National committee are invited to comment on the changes of the title from

“Part 6: Video content analytics – Performance testing and grading”
to

“Part 6: Performance testing and grading of real-time intelligent video content analysis devices & systems”.

Copyright © 2019 International Electrotechnical Commission, IEC. All rights reserved. It is permitted to download this electronic file, to make a copy and to print out the content for the sole purpose of preparing National Committee positions. You may not copy or "mirror" the file or printed version of the document, or any part of it, for any other purpose without permission in writing from IEC.

Table of Contents

1	Scope	10
2	Normative references	10
3	Terms, definitions and abbreviations	11
3.1	Terms and Definitions.....	11
3.2	Abbreviations	17
4	Overview standardized VCA Test procedure (normative).....	19
5	Description of VCA working principles and specifications (informative).....	20
5.1	General	20
5.2	The Application of this standard by Industry	21
5.2.1	Industry Users.....	21
5.2.2	What can industry participants do with this standard	21
5.2.3	Critical Operating Parameters	21
5.2.4	Video Input Quality.....	21
5.2.5	Testing Criteria	21
5.2.6	Table of Environmental Scenarios	21
5.2.7	Degrees of Difficulty (DoD).....	22
5.2.8	Testing setup configuration (no Ground Truth)	22
5.2.9	Testing setup configuration with Ground Truth	23
5.3	Categories of testing	23
5.3.2	Grading Criteria Levels	23
5.4	Working principle of Performance Test & Grading System	23
5.5	Video Format and System Requirements	24
5.5.2	Event Detection Scenarios	24
5.5.3	Standardised Classification for VCA Certification Tests	25
5.5.4	Key components that this standard contains;.....	25
5.6	Performance Testing Methodology	26
5.7	Grading Methodology	26
5.8	Criteria to be considered when using this standard.....	26
5.8.1	Light Requirements	26
5.8.2	Cameras	26
5.8.3	Recorded Footage.....	27
5.8.4	Normalised Input Video Signal	27
5.9	Most common VCA test scenarios with mandatory and applicable predefined test-sequences	27
6	VCA Performance Requirements and Measurements (normative).....	28
6.1	Moving object detection	28
6.2	Unattended object detection	28
6.3	Object removal detection	28
6.4	Wrong Direction motion detection	28
6.5	Moving object detection.....	29
6.6	Unattended object detection	29
6.7	Object removal detection	29
6.8	Line Crossing detection	29
6.9	Intrusion detection	29
6.10	Wrong Direction motion detection	30
6.11	Loitering detection	

6.12	Flow statistics.....	30
6.13	Density (congestion or overcrowding) detection.....	30
6.14	Object classification	30
7	XML Messaging Structure (normative).....	30
7.1	XML Indexing Schema (see Annex N).....	31
7.2	Event Detection Scenarios	31
7.3	Video description data output	33
7.4	Map File	33
7.5	List File	34
7.6	Ground Truth (GT).....	35
7.7	System Alarm (SA)	35
8	Test methods	37
8.1	Performance Testing Methodology	37
8.2	Grading Methodology	38
8.3	Criteria to be considered when using this standard.....	38
8.3.1	Light Requirements	38
8.3.2	Cameras	38
8.3.3	Recorded Footage.....	38
8.4	Test Sequence Procedures and Methods.....	38
8.5	Customer Requirement Tests (Conformity) Descriptions	39
8.5.1	Real-time intelligent video analysis test.....	39
8.5.2	Video description data output test	39
8.5.3	Alarm input and output test	39
8.5.4	Video output test.....	39
8.5.5	Video abnormality detection test	39
8.5.6	Device configuration test.....	39
8.5.7	Self-detection test	39
8.5.8	Factory default settings restoration test.....	39
8.5.9	Performance test.....	39
8.6	Automating the Performance Testing and Grading Process	40
8.6.1	Moving object detection.....	41
8.6.2	Unattended object detection.....	41
8.6.3	Object removal detection.....	41
8.6.4	Line Crossing detection.....	42
8.6.5	Intrusion detection.....	42
8.6.6	Wrong Direction motion detection.....	42
8.6.7	Loitering detection.....	43
8.6.8	Flow statistics	43
8.6.9	Density detection	43
8.6.10	Object classification	43
8.7	Sterile Zone Performance.....	44
8.8	Non Sterile Zones.....	44
8.9	Tolerance to Atmospheric Interference	44
8.10	Training, Test and Evaluation Video Footage datasets	44
8.11	Translating Specifications into XML – “Ground Truth” templates.....	44
8.12	Several Methods can be used to test performance as follows:	45
9	Typical Scenario Performance Criteria	45
9.1	Sterile Zone Performance.....	45
9.2	Non Sterile Zones.....	46

9.3	Tolerance to Atmospheric Interference	46
10	Object and Scenario Detection Messaging	46
10.1	XML Indexing Schema (see Annex N).....	46
11	VCA Object Detection Performance Specification (with examples)	47
11.1	Moving object detection	47
11.2	Unattended object detection	47
11.3	Object removal detection	47
11.4	Wrong Direction motion detection	47
11.5	Translating Specifications into XML – “Ground Truth” templates.....	48
12	System Grading Methodology.....	48
13	Automating the Performance Testing and Grading Process	48
14	Scenario Evaluation Criteria.....	48
14.1	Training, Test and Evaluation datasets	48
14.2	Typical Analysis Categorisation	49
14.2.2	Event Detection Scenarios	49
14.2.3	Standardized Classification for VCA Certification Tests	50
15	Data and XML Structure	50
15.1	Event Detection Scenarios	50
15.2	Video description data output	52
15.3	Alarm input and output	52
15.4	Video output	52
15.5	Video abnormality detection	52
15.6	Parameter configuration	52
15.7	Self-inspection	52
15.8	Restoration of factory default settings.....	52
16	Performance requirements (normative)	52
16.1	Moving object detection	52
16.2	Unattended object detection	53
16.3	Object removal detection	53
16.4	Line Crossing detection	53
16.5	Intrusion detection	53
16.6	Wrong Direction motion detection	53
16.7	Loitering detection	53
16.8	Count of objects	54
16.9	Crowd detection	54
16.10	Object classification	54
16.11	List File	55
16.12	Ground Truth (GT).....	56
16.13	System Event (Event XML File)	56
16.14	Test methods	58
16.15	Test Sequence Procedures and Methods.....	58
16.15.1	Performance test.....	58
16.15.2	Moving object detection.....	59
16.15.3	Unattended object detection.....	60
16.15.4	Object removal detection.....	60
16.15.5	Line Crossing detection.....	60
16.15.6	Intrusion detection.....	60
16.15.7	Wrong Direction (motion detection)	60
16.15.8	Loitering detection.....	60

16.15.9	Count (statistics)	60
16.15.10	Crowd detection	60
16.15.11	Object classification	60
17	System Evaluation (normative)	60
17.1	VCA Tests and Trials	60
17.2	Public Interaction Business Process Workflow – (Figure-7)	60
17.3	Sterile Zone Test Process (Workflow) Diagram Figure-8	61
17.4	Non-Sterile Zone Test Process (Workflow) Diagram Figure-9	62
17.5	Demonstration of the performance test process (results)	62
17.6	Techniques being used currently in country specific testing of VCA systems	63
17.7	Evaluation Score Threshold Examples	64
18	Evaluation Procedure (normative)	64
18.1	General	64
18.1.1	Event Detection	64
18.1.2	Object Tracking	65
19	Evaluation Method and Criteria (Korea)	65
19.1	Event detection functional description	65
19.2	Evaluation criteria	65
20	Automation tools for Grading (normative)	66
20.1	Table-I to homogenise the different video signal types	66
20.2	Test Setup Configuration	67
20.2.1	Source Stream	67
20.2.2	System under test accepts Source Stream	67
20.2.3	Results in XML output	67
20.2.4	Results - Evaluation Options	67
20.2.5	Important Prior Work	67
20.2.6	Configuration Considerations	67
20.3	Automated Spreadsheet Table-R to tabulate results and deliver grading	69
20.4	Method of Automating the Performance Testing and Grading Process	69
20.5	Grading Table	69
20.6	The Grading Calculation Table	70
21	Bibliography	71
Annex A	Real-Time Video Analysis Function Rules (Informative)	72
A.1	Description for Real-time Intelligent Video Analysis Function Rules	72
	Function rule data	72
	A.2 Rule capability set data	79
Annex B	Definitions of Data Types and the Coordinate System	83
	B.1 Description for data types	83
	B.2 Definitions of coordinates and the coordinate system	83
Annex C		84
C.1	Supplementary Description for Video Description Data	84
	C.1 Description for video description data type	84
	C.2 XML packaging example of video description data	84
Annex D	Real-time Intelligent Video Analysis Device (Video Footage List)	91
Annex E		112
E.1	Event Detection & Object Tracking System Evaluation	112
Annex F		113
F.1	FAQs	113

Annex G Scenario Definitions	114
G.1 Scenario Definition	114
G.2 Doorway Surveillance Scenario Definition.....	115
Annex H Performance Evaluation By Recording Behaviour Method	117
H.1 Outline	117
H.2 Korean Composition and scope	117
H.3 Ground Truth.....	117
H.4 Performance evaluation system.....	117
H.5 Video transfer device.....	117
H.6 Intelligent video analysis system	118
H.7 Performance evaluation device.....	118
H.8 Method of writing system alarms	118
Outline 118	
XML schema	118
Composition of system alarms.....	119
Annex I.....	122
I.1 List of CORE VCA capability.....	122
I.1.1 “Object Presence” in an area.....	122
I.1.2 “Object Stopped Moving”	122
I.1.3 “Object Started Moving”	122
I.1.4 “Object Appears”	122
I.1.5 “Object Disappears”	122
I.1.6 “Object Direction of movement”	122
I.1.7 “Object Speed”	122
I.1.8 “Target Lock”	122
I.1.9 “Object Wrong Direction”	122
I.1.10 “Object Removed”	122
I.1.11 “Object left behind”.....	122
I.1.12 “Object Speed”	123
I.1.13 “Object Classification”	123
Annex J.....	124
J.1 List of Complex VCA Capabilities	124
J.1.1 “Abandoned baggage detection”.....	124
J.1.2 Parked vehicle detection	124
J.1.3 “Sterile zone monitoring”	124
J.1.4 Perimeter Protection	124
J.1.5 Intrusion Detection	124
J.1.6 Unusual Events	124
J.1.7 Loitering.....	124
J.1.8 Person Down (Trip and Fall).....	124
J.1.9 Owner Tracking (Triggered by left luggage).....	124
J.1.10 Terrorist Attack (using Vehicle)	124
Annex K	125
K.1 XML Messaging.....	125
Annex L.....	128
L.1 Sensor Technology.....	128
Annex M.....	130
M.1 Video Standards (list).....	130
Communications Standards (list).....	130

Electrical Standards (list)	130
Software standards (list)	131
Embedded Hardware standards (list)	131
Ethernet Networking Standards	131
Safety Standards	131
Ethical and Regulatory Standards	131
Annex N Video Footage Playlist List	132
N.1 Video Footage Library and Playlist for use in Testing	132
Annex O XML Message Creation Tools	138
O.1 VCA Testing, Ground Truth, and XML Messaging Format Creation Tools	138
Annex P Degree of Difficulty Explanation	139
P.1 Degree of Difficulty Levels	139
Internal	139
External	139
Night	139
Cloudy (poor light and or high shadows)	139
Shake (Vibration and Wind shake)	139
Obscuration	140
Generate Alarm	140
Object Tracking	140
Annex Q Scenario Definitions	141
Q.1 Scenario Definitions	141
Sterile Zone Scenario Definition	142
Multiple Camera Tracking Scenario Definition	143
Annex R New Technology Explanations (iLIDS)	147
R.1 New Technologies Scenario Definitions	147
Annex S Encapsulation of Video Description Data	155
S.1 Encapsulation of Video Description Data Tables	155
Table 3 Encapsulation Format for Video Description Data	155
Table 4 Encapsulation Format for Object Information	156
Table 5 Encapsulation Format for Event Detection Information	158
Table 6 Encapsulation Format for the Rule Information of Unattended Object Detection/Object Removal Detection	159
Table 7 Encapsulation Format for the Rule Information of Line Crossing Detection	159
Table 8 Encapsulation Format for the Rule Information of Intrusion Detection	160
Table 9 Encapsulation Format for the Rule Information of Wrong Direction Motion Detection	160
Table 10 Encapsulation Format for the Rule Information of Loitering Detection	161
Table 11 Encapsulation Format for Flow Statistics Information	161
Table 12 Encapsulation Format for Density Detection Information	161
Table 13 Encapsulation Format for Video Abnormality Detection Information	162
Annex T Explanation of Object Tracking (iLIDS)	163
T.1 Object Tracking	163
Annex U Depth of Field Considerations	166
U.1 Depth of Field Considerations in VCA Performance	166
Annex V Using LED Monitor as Test Target	168
V.1 Using 2D Screen as VCA target	168
Annex W Aperture – Sensitivity – DoF & FOV	169

W.1 How Aperture Effects Sensitivity, Depth of Field & Field of View..... 169

Annex X BUSAN IEC Meeting Presentation 172

X.1 Demonstration Setup at BUSAN meeting 172

X.2 Purpose of Demonstration of Framework at IEC Busan meeting 172

X.3 Creating Video Footage for camera with embedded Video Content Analytics..... 173

X.4 Referencing the test Framework 174

X.4.1 The Video Source 174

INTERNETIONAL ELECTROTECHNICAL COMMISSION

VIDEO SURVEILLANCE SYSTEMS FOR USE IN SECURITY APPLICATIONS –**Part 6: Performance testing and grading
of real-time intelligent video content analysis devices & systems**

INTRODUCTION

The IEC Technical Committee 79 in charge of alarm and electronic security systems together with many governmental organisations, test houses and equipment manufacturers has defined a common framework for video surveillance transmission in order to achieve interoperability between products

The IEC 62676 series of standards on video surveillance system is divided into 6 independent parts:

Part 1: System requirements

Part 2: Video transmission protocols

Part 3: Analogue and digital video interfaces

Part 4: Application guidelines

Part 5: Data specifications and image quality performance for camera devices

Part 6: As above (Performance testing and trading of real-time intelligent video).

Each part offers its own clauses for the scope, normative references, definitions and requirements

The purpose of this part of IEC 62676 is to specify representation and measuring methods of performance values to be described in materials such as instruction manuals, brochures and specifications of Video Surveillance Camera equipment, and provide the convenience for users, installers, integrators and maintenance companies, etc.

VIDEO SURVEILLANCE SYSTEMS FOR USE IN SECURITY APPLICATIONS –

Part 6: Performance testing and grading of real-time intelligent video content analysis devices & systems

Scope

This Standard specifies the functions, performance, interfaces, environmental adaptability, test methods, performance evaluation and grading rules of real-time intelligent video analysis in surveillance systems.

This Standard applies to live and forensic, real-time intelligent video analysis devices and systems in video surveillance.

The standard is centred on testing performance and grading of device functionality which enables;

- **Core Capability:** Detection of specific "object activity" such as "stopping", "starting", "direction of movement", etc.
Examples and current expanded list in **Annex H**
- **Complex Capability:** Detection of "scenarios" which are based on combinations of object activity such as "loitering", "perimeter intrusion detection", "person down", "tail gating", "intrusion", "abandoned object detection", Explosion, Fire, Flood, Potential Terrorist attack using vehicle, Owner of abandoned bag, etc.
Examples of current scenarios are listed and described in **Annex I**

Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references only, the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC 61146-1: *Video cameras (PAL/SECAM/NTSC) - Methods of measurement - Part 1: Non- broadcast single-sensor cameras*

IEC 62676-1-1: *Video surveillance systems for use in security applications – Part 1-1: System Requirements*

IEC 62676-1-2: *Video surveillance systems for use in security applications – Part 1-2: General Video Transmission Requirements*

IEC 62676-2-1: *Video surveillance systems for use in security applications – Part 2-1: Video Transmission Protocols – General Requirements*

IEC 62676-2-2: *Video surveillance systems for use in security applications – Part 2-2: Video Transmission Protocols – IP interoperability implementation based on HTTP and REST services*

IEC 62676-2-3: *Video surveillance systems for use in security applications – Part 2-3: Video Transmission Protocols – IP interoperability implementation based on web Services*

IEC 62676-3: *Video surveillance systems for use in security applications– Part 3: Analogue and digital video interfaces*

IEC 62676-4: *Video surveillance systems for use in security applications– Part 4: Application guidelines*

IEC 62676-5:2018 *Video surveillance systems for use in security applications - Part 5: Data specifications and image quality performance for camera devices*

ISO 12233 Photography — Electronic still picture imaging — Resolution and spatial frequency responses 84
IEC CDV 62676-5  IEC 2016
ISO 14524 Photography — Electronic still picture cameras — Methods for measuring opto-electronic
conversion functions (OECFs)
ISO 15739 Photography — Electronic still-picture imaging — Noise measurements
ISO 17850 Photography — Digital cameras — Geometric distortion (GD) measurements
ISO 21550 Photography — Electronic scanners for photographic images — Dynamic range measurements
ISO/AWI 19093 Photography — Digital cameras — Measuring low light performance
ISO 18844/CD Photography — Electronic still picture imaging — Flare measurement techniques for digital
camera systems
ITU-R Recommendation BT.601, Studio encoding parameters of digital television for standard 4:3 and
wide-screen 16:9 aspect ratios
ITU-R Recommendation BT.709, Parameter values for the HDTV standards for production and
international programme exchanges
IEC 62471 Photo biological safety of lamps and lamp systems
IEC 60825-1: Safety of laser products — Part 1: Equipment classification and requirements
Refer to Annex R – for other Reference Standards

Terms, definitions and abbreviations

3.1 Terms and Definitions

3.1.1 Surveillance video

Information used for security monitoring and is composed of multi-frame pictures taken by
cameras, viewed live or stored on digital video storage technologies.

3.1.2 Acceptance Quality Limit (AQL)

Quality of “image” is determined by key measurable parameters such as “resolution”, “grey
scale depth”, “depth of field”, colour correctness.

3.1.3 Frames per Second (FPS)

The rate of capture of images by a camera device depends on performance parameters of
device and is measured in “Frames” per Second.

3.1.4 Motion Joint Photographic Expert Group (MJPEG)

A standards workgroup focused on creating and maintaining the MJPEG standard.

3.1.5 Reject Quality Level (RQL)

Image deterioration threshold established to ensure Video Analytics functions.

3.1.6 Software Development Kit (SDK)

Information provided to facilitate “interfacing” two systems using data and protocols to
establish two way communications between them.

3.1.7 Surveillance Video and Audio Coding (SVAC)

Hardware and Firmware embedded in a device to compress video and audio to meet
guidelines in the adopted standard.

137 **3.1.8 Objects**

138 Pedestrians, People, Vehicles, Animals, etc.

139 **3.1.9 Event**

140 An “event” notification occurs when a predefined “object activity” or “scenario” of interest is
141 detected and highlighted to the user and system alike.

142 **3.1.10 Alarm**

143 An “Alarm” notification occurs when an “object activity” or “scenario”, or “system malfunction”,
144 predefined as a security risk, is detected and annunciated to the operator.

145 **3.1.11 Object size**

146 The object size is defined by the width (horizontal pixel value) and height (vertical pixel value)
147 of the minimum bounding rectangle of the object in the video.

148 **3.1.12 Object position**

149 The centre coordinates of the defined object in the image; taken from the upper left corner of
150 the video, with the positive x-axis horizontal and the positive y-axis vertical.

151 **3.1.13 Real-time intelligent video analysis**

152 Real-time intelligent video analysis as opposed to retrospective intelligent video analysis, is
153 when the analysis occurs at the same time as the image is being captured, automatically
154 detecting the object or event of interest, then outputting the analysis result in text, picture or
155 video.

156 **3.1.14 Real-time intelligent video analysis device**

157 Devices that enables/facilitates real-time intelligent video analysis occurring.

158 **3.1.15 Video description data**

159 Data output from the real-time intelligent video analysis devices, describe the objects and
160 events in the video.

161 **3.1.16 Moving object detection**

162 An object is detected and an event created when it is moving within a predefined area of the
163 image.

164 **3.1.17 Unattended object detection**

165 An object is detected and an event created when the object has not moved after a
166 predetermined time within a predefined area of the image.

167 **3.1.18 Object removal detection**

168 An object is detected and an event created when the object has been removed for longer
169 than a predetermined time within a predefined area of the image.

170 **3.1.19 Line Crossing detection**

171 A Line Crossing is a virtual fence line that has been created in the camera’s field of view, that
172 when an object traverses the line in a specific or multiple directions creates an event.

173 **3.1.20 Intrusion detection**

174 A physical input from an external device creates and event when an incident occurs
175 triggering the camera or server to react.

176 **3.1.21 Wrong Direction (Retrograde) motion detection**

177 The detection of an object traversing an area or crossing a line in a direction that is not
178 deemed normal will determine an event.

3.1.22 Loitering detection

An object is detected and an event created when it has been moving in a predetermined area for greater than a predetermined time.

3.1.23 Count statistics

The number of objects counted in specified periods of time traversing lines or areas in specific directions.

3.1.24 Crowd or Congestion detection

An event is created when the count of classified objects (e.g. people or vehicles) exceeds the predefined threshold or density that was set - an example being "more than 3-people in front of an ATM machine".

3.1.25 Object classification

An object is detected and an event created when different categories of objects are identified, for example the detection of a car will create an event while the detection of people will not.

3.1.26 True Positive (TP)

An event is created when an object that should be detected appears in the field of view and the real-time intelligent video analysis device outputs the correct detection result.

3.1.27 False Positive (FP)

An event is created when an object that should be detected does not appear in the field of view and the real-time intelligent video analysis device outputs the correct detection result.

3.1.28 False Negative (FN)

The object or event that should be detected appears in the video, but the real-time intelligent video analysis device does not output the detection result.

3.1.29 Detected object

A detected object is one when the real-time intelligent video analysis device detects a time/item by the system and recognizes it as an object.

3.1.30 Reference data

Objects, events and other content in the video, that is either manually or artificially tagged and stored as "reference data" to which all other data is checked against.

3.1.31 Reference object

The object defined in reference data.

3.1.32 Object matching ratio

The degree of closeness between detected object and reference object displayed as a ratio.

Object matching ratio=

$$\frac{\text{No. of pixels in the overlapped area of the detected object and the reference object} \times 2}{\text{Number of pixels of the reference object} + \text{Number of pixels of the detected object}}$$

3.1.33 True positive rate

The frequency of correct objects or events outputted by the real-time intelligent video analysis device and the number of objects or events in the video that should be detected by the device displayed as a percentage.

True positive rate=

$$\frac{\text{Number of true positive results} \times 100\%}{\text{Number of false negative results} + \text{Number of true positive results}}$$

3.1.34 False positive rate

The frequency of mistaken objects or events measured against the total number objects or events outputted by the real-time intelligent video analysis device displayed as a percentage.

224 **False positive rate=**
 225
$$\frac{(\text{Number of false positive results}) \times 100\%}{\text{Number of false positive results} + \text{Number of true positive results}}$$

 226

227 **3.1.35 Classification accuracy**

228 The number of correct object classifications made by the real-time intelligent video analysis
 229 device against the number of objects displayed written as a percentage
 230

231 **Classification accuracy =**
 232
$$\frac{\text{Number of correctly-classified objects} \times 100\%}{\text{Number of correctly classified objects} + \text{Number of incorrectly classified objects}}$$

 233
 234

235 **3.1.36 Detected object or event**

236 The real-time intelligent video analysis device shall action a preconfigured response such as
 237 show video and outputs "meta-data" with a text base description of the event and parameters
 238 in a chosen protocol format e.g. - XML.

239 **3.1.37 Sterile Zone**

240 A sterile zone is an area where the presence of any human being, animal, vehicle or object is
 241 prohibited. If an object enters the zone, it is considered a high-level breach, to this closely
 242 monitored area and the reaction to a breach is should trigger an emergency response.

243 **3.1.38 Non-Sterile Zone**

244 Non-sterile zones are areas where objects entering and leaving are to be expected, which
 245 could be with basic restrictions like direction of movement or number of people who can be
 246 present at that location.

247 **3.1.39 Obscuration**

248 In any normal scene, "movement" of some kind will occur. People moving in front of and
 249 behind objects that are required to be under constant surveillance, e.g. an "object" left in a
 250 crowded scene will be obscured by moving people.
 251 The ability of the Video Analytic System to tolerate "obscuration", or the target obscured by
 252 another object for a period, is a performance function that the Video Analytic System is
 253 expected to deliver.
 254 So a video analytic system, which can perform at 50% obscuration, would be considered of
 255 significant value to a security surveillance operations team.

256 **3.1.40 "The Presence" of object when it enters an area**

257 The presence of a moving object (person or vehicle) in an area is when the object is given a
 258 "tracking" number and is constantly tracked even at 50% obscuration within that area and will
 259 maintain the same object tracking number while in the scene.

260 **3.1.41 Stop Moving**

261 Tracked object will have many key states, stop, start, random movement and fast (speed)
 262 movement but the system will send out a notification if the object stops for a predefined
 263 period.

264 **3.1.42 Start Moving**

265 Tracked object will have many key states, stop, start, random movement and fast (speed)
 266 movement; but when an object stops for predetermined period of time and starts moving again,
 267 a notification is sent out.

268 **3.1.43 Appears**

269 An object appears, when the thing/item becomes visible or detectable, especially without
 270 apparent cause such as from behind a fixture that obscured the object from the VCA system
 271 when it first set itself up. (Such as a person exiting a vehicle in a scene or emerging from
 272 behind a fixture which is already part of the field of view of the VCA camera).

273 3.1.44 Disappears

274 An object disappears, when an object becomes invisible or is removed from in between the
275 field of view, especially behind a fixture which is already part of the scene when the VCA
276 system first set itself up. (E.g. a person getting into a vehicle or getting behind a fixture which
277 is already in the VCA systems field of view).

278 3.1.45 Direction of movement

279 Direction of movement has many trajectories that can be tracked; up, down, forwards,
280 backwards, right and left.

281 3.1.46 Alarm

282 An alarm is an audible and/or visual notification of a change in condition usually identifying
283 an event.

284 3.1.47 Object Tracking

285 An object entering a scene (either manually or automatically) is given a unique ID that
286 identifies the object in the scene and this allows that object to be tracked through the scene.
287 Tracking information is used to “stitch” the objects movements across each camera’s field of
288 view, across multiple cameras and areas.

289 3.1.48 Speed

290 The movement of an object through a scene (the distance) is analysed in time and then
291 categorised as “speeding” when the movement breaks a pre-defined threshold setup for the
292 object. This helps differentiate between tracked objects and provides a parameter on the
293 object, which can be reported as unique to the objects behaviour / pattern of movement.

294 3.1.49 Classification

295 Object classifications such as “person”, “vehicle”, “truck”, “cycle etc. are “preconfigured” (i.e.
296 created using size, shape, aspect ratio, etc.) to enable the “video content algorithm” to “match”
297 objects entering the scene to these preconfigured criteria so they then become a valid “target”
298 for the system to track.

299 3.1.50 Target

300 An object, which has been detected by the video content analysis algorithm as meeting the
301 criteria or classification of a “target” to be tracked. For example an object enters a scene and
302 fits the object the classification of a “vehicle” or “person” the system will begin the tracking
303 process by giving that object a fixed ID number.

304 3.1.51 Object Stopped

305 When a target that is being tracked stops for a preconfigured period.

306 3.1.52 Object Started

307 When a target being tracked is marked as stationary then after a preconfigured period and
308 starts to move.

309 3.1.53 Object Direction

310 The detection of a targets movement in a predefined direction, taking into account the
311 object’s “width” and angle of entry etc.

312 3.1.54 Object Wrong Direction

313 The detection of a targets movement counter to a predefined direction, taking into account
314 the object’s “width” and angle of entry etc.

315 3.1.55 Object Removed

316 An object that has been present/static in a scene is no longer there and the absence duration
317 is greater than a predefined period.

3.1.56 Object Presence

An object that has not been present/static in a scene is now there and the appearance duration is greater than a predefined period.

3.1.57 Object Tracking Number

Objects entering the scene that match the preconfigured classification criteria become a "target" and are given a valid "tracking" number by the system.

3.1.58 Object left behind

Objects entering the scene that match the preconfigured classification criteria becomes a valid "target" and are given a "tracking" number by the system if they have remained static for a predetermined period.

3.1.59 Object Speed

Tracked objects (Targets) are tracked in real-time and if their movement (change in position) is greater than or less than a predetermined value then an event is triggered.

3.1.60 Image Stabilisation

CCTV cameras are prone to vibration and shock when strong winds, vehicle movement, storms, earthquakes etc. occur. Most of them have a built in "Image Stabilization" algorithm (function) to compensate for this movement to ensure the "video analytic engine" gets a steady image to analyse.

3.1.61 Tolerance of light changes

CCTV cameras are designed to compensate for light changes, day-night setup requirements, colour balance, highlight suppression, etc. To ensure that the "video analytic engine" is presented with a stable image, to prevent erroneous analysis and detection. The camera shall also cope with any sudden light changes such as flaring caused by a slow auto-iris response.

3.1.62 Privacy Masking

Privacy masking is the ability of the camera or system to hide/mask predetermined areas from the view of the operator.

3.1.63 Aspect "objects in the distance" error correction or compensation

The setup should allow the "detection algorithm" to compensate for "aspect" i.e. distance from camera object is smaller and hence at 1m the algorithm should look for object at size x and at 5m distance from camera the object detected size should be adjusted for that distance. Note: tools that help easily set up compensation are generally provided with "mannequins" to help with WSYWIG type setup.

3.1.64 Tamper Resistance

Changes to camera position, image quality (focusing change, image obscuration shall result in a tamper event alarm notification.

3.1.65 Video resolutions

Sources of video such as cameras & encoders generate Analogue and Digital video to specified resolutions. Analogue systems deliver CIF, 4CIF (720H) or 960TVLH (using encoding techniques such as CVI, TVI, etc.). IP Digital Video from Megapixel, HD (1080 x1920) to 4K, 8K & odd "multi megapixel" imager densities. Encoded video at higher resolutions should also be considered in the image processing options.

3.1.66 Video Content Analysis

The process of analysing the video image content within a field of view, to detect recognisable patterns, objects or activity or patterns of activity over time.

3.1.67 Machine Learning

The process of enabling the Video Content Analysis device to autonomously "analyse the scene", learn from specific changes that meet predefined criteria and use the "learning" to improve detection performance.

3.1.68 Artificial Intelligence

The behaviour of a system that is capable of improving its functional performance using historical training data or data generated by autonomously and continuously “analysing the scene” to update itself with corrected observations and analysis outcomes to improve its performance over time to meet the pre-set detection objectives.

3.1.69 Neural Network

The high throughput “analytic processing engine” embedded in the Video Analytic devices to speed up and improve content analysis.

3.1.70 Autonomous devices or systems

Devices or systems given a surveillance or response objective and perform this task independently. For example, automated car parks, where the system reads the vehicle number plate when the vehicle enters the area and compares this to previous stored authorised data when the vehicle or object leaves the area.

3.1.71 Surveillance Robots

Autonomous robotic devices programmed to conduct a specific routine task such as travelling around a perimeter or attending a rear entrance whilst performing a series of surveillance tasks. (Currently used in shopping centres in the US, also built as UAVs and Drones.)

3.1.72 Multi-camera Tracking

Multiple camera views that are virtually stitched together to create a single “virtualised image” for the video analytics system to track the path of travel of an object through the extended scene and also provide “business intelligence” of that large scene for security and safety purposes.

3.1.73 Integrated Analytic System

Multiple video analytic devices or systems, working together, each sending data that is gathered independently, to then conduct a more accurate verification or analysis to generate a speedier and more accurate response directly to the first responder.

3.2 Abbreviations

CCD Charge Coupled Device

CIF Common Intermediate Format

CIE Commission Internationale de l'Éclairage (International Commission on Illumination)

CMOS Complementary Metal Oxide Semiconductor

Data Base Decibel

DHCP Dynamic Host Configuration Protocol

DNS Domain Name System

DoD Degree of Difficulty

DVI Digital Visual Interface

FA False Alarm Rate

FPS Frames Per Second (frame rate)

FTP File Transfer Protocol

GB Giga Bytes

GPU Graphical Processing Unit

HD High Definition (1280 x 720, 1920 x 1080)

408	HDMI High Definition Multimedia Interface
409	HTTP Hypertext Transfer Protocol
410	IEEE Institute of Electrical and Electronics Engineers
411	IP Internet Protocol
412	IR Infra-Red
413	ISO International Standards Organization
414	JPEG Joint Photographic Experts Group
415	kB Kilo Bytes
416	kb/s Kilobits per second
417	LED Light Emitting Diode
418	MB Mega Bytes
419	Mb/s Megabits per second
420	MPEG Moving Picture Experts Group
421	ND Neutral Density
422	NTP Network Time Protocol
423	NTSC National Television System Committee
424	OECF Opto-Electronic Conversion Function
425	PAL Phase Alternating Line
426	PG Performance Grade
427	SD Standard Definition
428	SDI Serial Digital Interface
429	SFR Spatial Frequency Response
430	SHA Secure Hash Algorithm
431	S/N Signal to Noise
432	SMTP Simple Mail Transfer Protocol
433	SNR Signal to Noise Ratio
434	SSL Secure Socket Layer
435	TCP/IP Transmission Control Protocol / Internet Protocol
436	UDP User Datagram Protocol
437	VGA Video Graphic Array
438	UAV Unmanned Ariel Vehicle or Underwater autonomous vehicle
439	VCA Video Content Analytics
440	VCAC VCA Capability
441	VCATP VCA Test Procedure.

Overview standardized VCA Test procedure (normative)

This specification standardizes:

- VCA test procedures of test-houses, -labs and performance-benchmarks of end-users, vendors, etc.
- Metadata-Interface for the VCAC system such as devices, software applications, services or other implementations
- any VCA test-scenario in a framework
- an automated process to compare the generated VCAC metadata with ground truth
- the most common and important test scenarios using a set of predefined test-sequences and footage.

This results in the capability to grade VCAC performance. This enables the comparison and ranking of similar test-scenarios of the different test-houses including the application of the degree of difficulty

- the declaration of performance

This standard is applicable to

- VCAC manufactures
- test-houses and -labs for existing or upcoming VCAC test-procedures and certifications, as an overall framework.
- installers, integrators

This standard specifies following mandatory requirements:

1. The VCAC system shall fulfil the following requirements:

- a. support the standardized **Video input stream format** for VCA test purposes
- b. classify objects and generate alerts using the standardized metadata format (*chap.7,11,15*)

This metadata shall contain information related to the scenario and events,

This metadata stream can be parsed into a table of results, allowing the comparison with ground truth data

This collected data can use well defined and specified formulae to generate the performance grading

- c. capable of VCA **Configuration Export and Import** to ensure that the test parameters are logged to allow the reproduction of results
- d. offer the standardized datasheet

2. The standardized VCA test report shall be composed of the VCAC config, parsed metadata, and detailed report

3. The VCA test- and certification-house shall offer a **Declaration** of Performance DoP including an overall grade in the following manner:

e. Specify one of the following VCA test fields:

- i. **Performance Test and grading of the VCAC system** used and established by test houses resulting in overall performance grades e.g. for very specialized scenarios
- ii. **Compliance to specifications** for installed systems (already in operation)
- iii. **Conformance** to available test definition e.g. after repairs and calibrations, offered by customers

f. Include the details listed in this specification such as:

i. *Test results (false alarm rate in % FA, detection probability DP in %,...) of the requested scenario such as perimeter detection for*

1. **lab-specific test sequences and playlists**

2. **standardized test-sequences and playlists defined by this specification for the same scenario class, if scenario and test sequence is defined by this standard**

ii. **resulting overall performance grade PG** for this test-scenario, normalized with the degree of difficulty DoD associated with the used test scenario and sequences

iii. *reference to this specification IEC 62676-6, the tested-scenario such as 'perimeter protection', the used public test-sequence via its internet link www.iec.vca-testseq.com/perimeterprotectionPP1A.mp4*

iv. **Quadrant and name of tested scenario with Degree of Difficulty DoD**

v. **List of disturbance types** applied in test video such as light changes, rain, birds, insects, etc.

vi. **Public Internet Link to the video footage training set of the tested scenario**

vii. **Other details listed in this specification**

•

Description of VCA working principles and specifications (informative)

5.1 General

The standard provides guidance on performance testing and grading of:

- a) Devices or systems that autonomously analyse a scene for specific targeted object activity are listed in Annex L, in a live camera video surveillance ~~scene~~ or on forensic analysis on recorded footage on a scene. When observed activity matches that which is selected during the test, a notification will be generated with the necessary descriptive data in an XML format.
- b) Devices or systems that autonomously analyse a scene for complex capability scenarios, "behaviours" or "activity", are listed in Annex M, in a live camera video surveillance or on forensic analysis of recorded footage on a scene. When observed activity matches that which is selected during the test, a notification will be generated with the necessary descriptive data in an XML format.

- c) Meeting the IEC compliant XML messaging (Metadata) standard, IEC62676-4, that provides the “descriptors” related to “detection” of “behaviours” or “activity” listed in Annex L & M that meet the management and GRADING requirements.
- d) Methods adopted to develop “ground truth” metadata (XML) to confirm performance outcomes as per examples are listed in Annex N.

5.2 The Application of this standard by Industry

5.2.1 Industry Users

The standard is intended for use by industry at different levels such as;

1. Manufacturers, to design and tests in-house to their specification.
2. Test and Certification Centres who can test to ensure that a customer’s or manufacturer’s product range operates to the functional specification listed in this standard and will be able to certify it meets this standard with a grading level that reflects the performance of the products during the specified tests.
3. System Integrators to test field installed systems after installation, commissioning or repair of systems by field technical teams.
4. System Integrators to verify customer functional requirements (specific scenarios or object detection requirements).
5. Acceptance Tests done by or for customers based on guidance provided in this specification.

5.2.2 What can industry participants do with this standard

The industry can test the performance of VCA devices and systems using live and recorded video footage created by different imaging Technologies such as:

- a) Visible light spectrum (400 to 700 nanometres)
- b) Near IR (700nano meters to 1micrometer)
- c) Thermal (8 to 14 micrometres)
- d) The industry can test performance of “Intelligent Video Analytic” systems that use processing technologies such as AI, ML & Neuralnet Technology, etc., that continuously enhances performance based on built-in “learning” and “reasoning” capability.

Examples of these technology deployments are shown in Annex O

5.2.3 Critical Operating Parameters

Video Content Analytics (VCA) technology and “Intelligent” Video Content Analysis technology (which uses ML, AI, Neuralnet Technology, etc. to enhance operational functionality) has a wide application base. This standard is focused on the application of Video Content Analysis as applied to safety and security.

5.2.4 Video Input Quality

The performance of any VCA device will depend on the quality of video input provided to it. Hence the Video Content and Signal Quality (VC&SQ) will need to comply with the relevant IEC Standards listed under normative references.

5.2.5 Testing Criteria

Performance Testing Metrics (PTM) will be established for different environmental scenarios. These scenarios will be derived from currently documented customer requirement and future aspirational criteria. Degree of Difficulty (DoD) levels will also be established as part of the performance testing criteria.

5.2.6 Table of Environmental Scenarios

The table below presents the Environmental Scenario schema for the Tests

5.2.7 Degrees of Difficulty (DoD)

The Degree of Difficulty measure is applied to video content analysis to confirm if the application of the technology is possible in various environmental and situational scenarios at the application site.

Table 1 Environmental Test Scenarios

Environmental Test Scenarios All tests are 24hr Tests (Day Night)								
	Sterile Zone (No "objects" in the scene)				Non Sterile Zone (Objects in the scene causing obscuration)			
	Indoor (No Lighting variations)		Outdoor (light variations)		Indoor (No light variation)		Outdoor (light variations)	
Names	Test	Degree of Difficulty	Test	Degree of Difficulty	Test	Degree of Difficulty	Test	Degree of Difficulty
See Annex L & M	Video footage list in Annex Q	DoD as in Flow chart in Annex XX	Video footage list in Annex Q	DoD as in Flow chart in Annex XX	Video footage list in Annex Q	DoD as in Flow chart in Annex XX	Video footage list in Annex Q	DoD as in Flow chart in Annex XX
(Example) Loitering	(Example) V001	(Example) Level 0	(Example) V002	(Example) Level 6	(Example) V007	(Example) Level7	(Example) V020	(Example) Level 10

Two main categories have been defined; Sterile Zone & Non Sterile Zone.

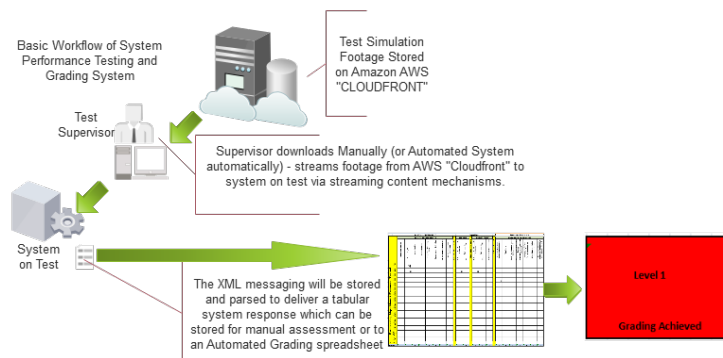
In each category there is consideration for the following key operational aspects;

1. Indoor activity (where light is considered to be at a constant level or changes predictably)
2. Outdoor activity (where light is a "variable", example clouds and "random" because it is unpredictable)
3. Degree of Difficulty (Conditions do vary at a site, for example rain, snow, shake, fog, etc)

Performance testing using this framework produces a "value" in almost all operating conditions.

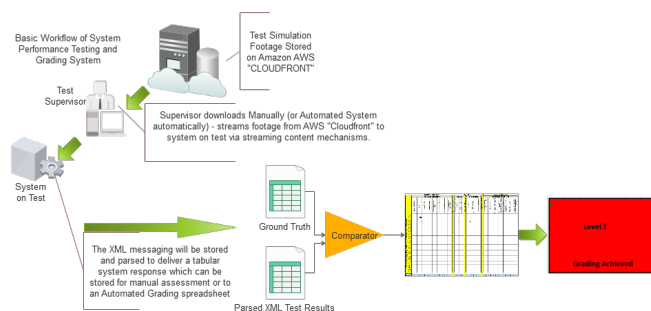
5.2.8 Testing setup configuration (no Ground Truth)

The diagram provides an overview of a test configuration using no ground truth.



5.2.9 Testing setup configuration with Ground Truth

The diagram provides an overview of a test configuration with ground truth



5.3 Categories of testing

The categories of testing will be;

5.3.1.1 Performance Testing

Performance testing will test for “Typical Functionality” as listed in Annex L & M.

5.3.1.2 Customer Requirement Testing

Customer Requirement are tests set to determine that a customer specification such as generating event alarms and/or functional criteria such as “running” are met. e.g.: speed of human object moving above 5km/hr.

5.3.1.3 Stress Testing

When the VCA device is put through endurance tests – e.g.– day/night and bad weather scenes, specific object detection, obscuration of objects and picture shake (degree of difficulty), which are included in the playlist video footage for this kind of tests – it will be termed “stress tests”. The report will present how the system performed throughout the footage playlist. Performance will be described as a percentage of the actual “objects” detected at different levels at which the “degree of difficulty” is applied as presented in the test footage.

5.3.1.4 Maintenance testing

It is recommended that the system should be maintained annually ~~due~~ to take into consideration potential changes in environmental conditions.

5.3.2 Grading Criteria Levels

Based on the Performance Testing Metrics (PTM) ~~chosen~~ and Degree of Difficulty (DoD) grading levels are applied. Grading levels have been established as shown in Table x and listed in Annex XX

5.4 Working principle of Performance Test & Grading System

The working principle of the device is shown in Figure 1 and shall include the following content:

Inputs: The device shall support inputs from one or more sources of video inputs; analogue, digital video signals, network digital video bit streams, audio data, sensor alarm data and control data.

Data processing shall include input data pre-processing module, intelligent analysis processing module, output data processing module and a device control module:

The input data pre-processing module performs a normalisation of input signals into the VCA system, making format conversion of analogue, non-network digital video signals and network digital video bit streams to signals compliant to IEC62676-5.

The video content analysis module performs real-time intelligent analysis of the normalised video stream and/or stored video footage stream and generates video description data in XML and response signals as required by the performance testing process.

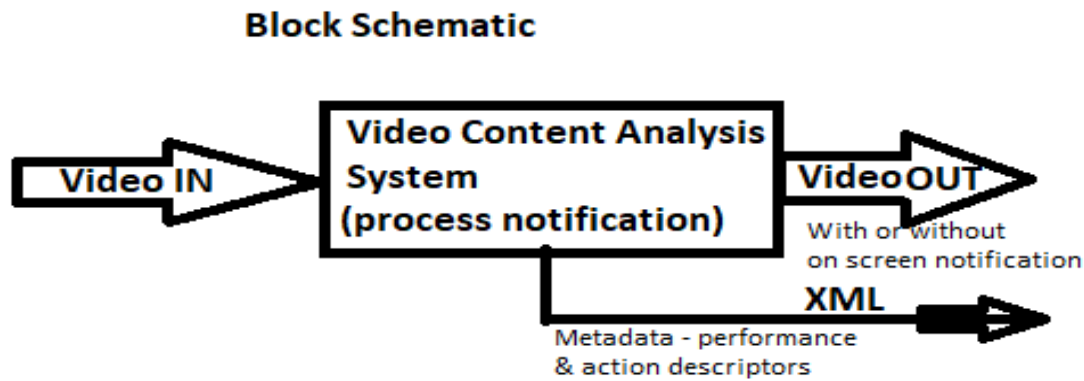


Figure 1 Working Principle of the Device

5.5 Video Format and System Requirements

The Event Simulation Video Footage should be in easily accessible formats (MJPEG, H.264, etc.) full frame or streaming file formats and accessible via real-time streaming protocol such as RTSP.

Typical Analysis Categorisation

5.5.1.1 C、Statistics character:

Counting statistics, Crowd detection (these outputs are statistical)

5.5.1.2 D、Quality character:

Video abnormality (the outputs are quality metrics)

The XML file includes hierarchical structure, while a table cannot. Especially for the target analysis, the detected target may have a subcategory. (So maybe we should use more than one table for database).

5.5.1.3 Condition Criteria:

An explanation and examples of this follows;

5.5.1.3.1 Degree of Difficulty Condition:

In order to test the performance of a VCA product, a “CONDITION GRADE” such as “Very difficult, Difficult, Ordinary, Easy” should be used.

5.5.1.3.2 Situation or Character Condition Examples:

While "CONDITION TYPES"("Day", "Night", "Sterile", "Non-Sterile", "Perimeter-only") can be used to describe the character of test sequence.

There are scenarios highlighted and explained in this standard as examples of where and how Video Content Analytics is applied. There are many more highlighted in Annex L & Annex M and more will be added as the Standard is revised with contributions from industry.

5.5.2 Event Detection Scenarios

Typical scenario examples are;

5.5.2.1 Abandoned baggage detection

Are alarm events that occur when an unattended bag is detected (example) on a platform of an underground station.

5.5.2.2 Parked vehicle detection

Are alarm events that occur when a suspiciously parked vehicle is detected in an urban setting.

5.5.2.3 Sterile zone monitoring

Are alarm events consisting of the presence of people in a sterile zone between two security fences.

5.5.2.4 Detection using Thermal Imaging & IR Illumination Technologies

Are alarms that can be generated using VCA in total darkness using cooled and un-cooled thermal imaging sensors and/or infrared illumination technologies.

Scenarios determined by customer inputs and VCA functionality can be created by integrating various basic detection algorithms (Annex H) together to form other complex capability scenarios detection functions (Annex I).

**5.5.3 Standardised Classification for VCA Certification Tests**

The certification tests should consider the following use case classifications:

- a) Observation and Application purpose (Environment such as Outdoor and Extreme Outdoor)
- b) Level of complexity and degree of difficulty (Considering the grade of weather and light changes, length of test dataset and the number of objects)
- c) Level of Object Detectability
- d) Level of Nuisance Alarm Rate



Each Event Detection video dataset could comprise of approximately 24 hours of 'sequences', recorded in different conditions; time of day, weather, background activity etc. managed as a "playlist".

The Thermal and IR technology video footage dataset could contain approximately 48 hours of sequences managed as a "playlist" for each imaging modality.

Some sequences are augmented by alarm events 'acted out' for the footage. The remaining 'non-alarm' sequences contain only a background level of alarm events.

In the public video footage training datasets, 'alarm' sequences containing many alarm events and are edited into 'clips' such that each alarm event acted out is rendered to a separate file managed as a "playlist". This makes it quicker for training dataset users to access footage of specific alarm events. The file naming convention is such that adjacent clips can easily be identified and concatenated should more pre- and post- event footage be needed e.g. for learning algorithms.

Each alarm sequence should contain at least 3 minutes of footage prior to the first scheduled alarm event so as to assist the learning systems in adapting to the conditions of the sequence

Video Content Analysis systems are generally required to report an Event or Alarm when a prescribed "activity" is detected in the scene.

Current applications of Video Content Analysis requires detection of "activity" associated with a-typical behaviour (unusual) of people and/or objects and may be analysed for non-security requirements such as "trend analysis".

Video Content Analysis Systems generate "data" in an XML format that contains a description of the scene, activity and detection categorisation. These elements are used to create more effective situational awareness.

5.5.4 Key components that this standard contains;

A list of core Video Content Analytics capability of interest to security and safety (Annex H)

A list of a few complex Video Content Analytic capabilities, from the many that can be constructed from core capability of interest to security and safety. (Annex I)

Standards based XML messaging by VCA systems generate usable actionable information which can be used for live control, historical and forensic analysis.

A method to “performance test” these VCA system will provide the industry with the tools to evaluate systems and verify if the system meets industry customer specified expectations.

A method to “grade” VCA systems to make the task, of selecting the right system for an application, easier.

5.6 Performance Testing Methodology

Events and Scenarios used in the performance and grading process, will be created using “video footage” generated from typical site activity suggested by industry contributors.

The video footage list is shown in Annex M

A testing methodology has been suggested to meet the following criteria:

- 1) The testing regime which will specify the type of footage to be used for the level of grading required.
- 2) The number of “false positives” (incorrect detections on the positive side), “false negatives” (incorrect detections on the negative side) and “correct detections” will be documented.
- 3) A performance evaluation formula will use the count numbers generated by the tests (in 2) to calculate, using several weighting factors and generate a “unique” performance level grading figure.
- 4) A report that incorporates methodology, performance tests and grading should be compiled to ensure users of the standard can be confident that the grading outcomes calculated is evidence based.

Testing methodology, formulae applied and reports generated are detailed in Annex XX

5.7 Grading Methodology

Systems that have passed have been through tests which evaluate the following:

- a) Individual Functionality
- b) Capability range and concurrency
- c) Performance under “degrees of difficulty”

Each of these by itself will establish if the grading can be classified as;

- a) Performance Grade compliant
- b) Conformance Compliant(to customer spec)
- c) Stress Complaint to prescribed levels (performance under said degree of difficult – for example requirement in applications under heavy wind and temperature variations)

5.8 Criteria to be considered when using this standard

Common requirements that will apply to the Performance Testing and Grading test environment include:

5.8.1 Light Requirements

Video Analytic Systems must be able to perform and be tested in all light and weather conditions, if required by the end user or customer.

5.8.2 Cameras

Video Surveillance cameras should deliver a usable video picture 24/7 (Ref: 62676-5), in all lighting conditions and should deliver a continuous “image stream” with little or no introduced interference. It is therefore incumbent on the manufacturer to test cameras and associated products (lenses, IR light etc.) to ensure the camera delivers the correct video levels to the video analytic system to enable it to perform to a satisfactorily level. The VCA system and industry

be tested when the inputs into the system are normalised and have no errors in the video signal.

5.8.3 Recorded Footage

Recorded footage of typical scenarios from a customer site or from footage archives of purpose created video clips, that mimic the exact site scenarios the system will be embedded into, should be used to conduct the tests. Such techniques will provide the user of this specification the opportunity to test the system using parameters recorded in test footage from the actual location of use.

5.8.4 Normalised Input Video Signal

The VCA system should only be tested when the inputs into the system are Normalised or checked against Video Standards using appropriate test equipment.

5.8.4.1 A、Target analysis:

Object detecting, object classification, moving object tracing are target analysis and the outputs are objects with properties, such as “bounding boxes”. Object “tracking number”, “Traces”, etc. become part of the evaluation and performance criteria.

5.8.4.2 B、Event analysis:

Line Crossing, Intrusion, Unattended object (abandonment), Object removal, Wrong Direction motion and Loitering (the output of these are alarm events)

5.9 Most common VCA test scenarios with mandatory and applicable predefined test-sequences

If a VCAC declares to support one of the following widely used detection scenarios the following listed and linked test scenario shall be applied and graded next to the optional lab specific ones:

- for outdoor Sterile Zones: Perimeter Intrusion Detection Test-sequence
<http://iec-vca-testseq.com/perimeterprotectionPP1A.mp4>

- for outdoor Sterile Zones: Perimeter Intrusion Detection with Thermal imaging
Test-sequence <http://iec-vca-testseq.com/ThermoperimeterprotectionThP1A.mp4>

- for indoor non-sterile zones: Wrong Way People Detection Test-sequence
<http://iec-vca-testseq.com/WrongWayWWP1A.mp4>,

- for indoor non-sterile zones: People Crowd Detection Test-sequence
<http://iec-vca-testseq.com/CrowdPC1A.mp4>,

- for indoor Sterile Zones: Indoor Intrusion Detection Test-sequence
<http://iec-vca-testseq.com/indoorIntrusionII1A.mp4>

- for indoor non-sterile zones: People Queue Detection Test-sequence
<http://iec-vca-testseq.com/People QueuePQ1A.mp4>,

- for indoor non-sterile zones: People Counting Test-sequence
<http://iec-vca-testseq.com/People CountingPC1A.mp4>,

- for outdoor non-sterile zones: Wrong Way Vehicle Detection Test-sequence
<http://iec-vca-testseq.com/WrongWayWWV1A.mp4>,

- for indoor non-sterile zones: Abandoned Object Detection in highly frequented scenes
Test-sequence <http://iec-vca-testseq.com/AbandonedObjAO1A.mp4>,

- for indoor non-sterile zones: Removed Object Detection Test-sequence
<http://iec-vca-testseq.com/RemovedObjRO1A.mp4>,

- for outdoor non-sterile zones: Parked Vehicle Detection Test-sequence
<http://iec-vca-testseq.com/ParkedVehiclePV1A.mp4>,

- for outdoor sterile zones: Harbor Waterfront Intrusion Detection Test-sequence
<http://iec-vca-testseq.com/MarinasMA1A.mp4>,

817 - for all outdoor: outdoor Tamper Detection Test-sequence
 818 <http://iec-vca-testseq.com/TamperIndoorTI1A.mp4>,

819 - for all indoor: indoor Tamper Detection Test-sequence
 820 <http://iec-vca-testseq.com/TamperOutdoorTO1A.mp4>,

821

822 The test and grading of this standardized test scenario and sequence allows the comparison of the
 823 lab-specific one in terms of DoD

824 **VCA Performance Requirements and Measurements (normative)**

825 VCA performance criteria is formulated and requested in two ways;

- 826 a) Using a written specification which will list out dimensional and temporal criteria
 827 required to be met.
- 828 b) Using typical scene footage from customer premises.

829 In both cases XML will be used to reconstruct or structure the temporal parametric description
 830 of the scene.

831 Several examples listed below provide detailed textual parametric descriptions which need to
 832 be translated into a common accepted XML message structure.

833 Review message construction methodology in Annex N.

834 A constructed XML message sample is shown in Annex N for simple moving object detection;

835 (*XML example – moving object detection – homogenised XML from iLIDS, China, Korea and ONVIF*)

836

837 **6.1 Moving object detection**

838 The requirements to characterise this detection are suggested below;

- 839 a. Shall be able to detect a single moving object with horizontal or vertical speed being in
 840 the range of 15 pixel/s~200 pixel/s and width and height being ≥ 16 pixel;
- 841 b. The true positive detection rate of a single object shall be $\geq 95\%$;
- 842 b) The false positive detection rate of a single object shall be $\leq 20\%$;
- 843 c) Shall be able to detect multiple objects simultaneously.

844 **6.2 Unattended object detection**

845 The requirements to characterise this detection are suggested below;

- 846 a. Shall be able to detect an object with width and height being both ≥ 32 pixel
- 847 b. The time to determine the stay of an object can be set in the range of 4s ~300s
- 848 b) The time to output the detection result shall be ≤ 1 s
- 849 c) The true positive detection rate shall be $\geq 90\%$
- 850 d) The false positive detection rate shall be $\leq 20\%$

851 **6.3 Object removal detection**

852 The requirements to characterise this detection are suggested below;

- 853 a. Shall be able to detect a removing object with width and height being both ≥ 32 pixel
- 854 b. The time to determine the removal of an object can be set in the range of 0 s~10 s
- 855 b) The time to output the detection result shall be ≤ 1 s
- 856 c) The true positive detection rate shall be $\geq 90\%$
- 857 d) The false positive detection rate shall be $\leq 20\%$

858 **6.4 Wrong Direction motion detection**

859 The requirements to characterise this detection are suggested below in the following table :

- a) Shall be able to detect a retrograding object with horizontal or vertical speed being in the range of 15 pixel/s~200 pixel/s, and width and height being both ≥ 32 pixel;
- b) The distance to determine Wrong Direction shall be settable in the range of 32 pixel~256 pixel;
- d) The time to output the detection result shall be ≤ 1 s;
- e) The true positive detection rate shall be $\geq 95\%$;
- f) The false positive detection rate shall be $\leq 10\%$.

6.5 Moving object detection

The requirements to characterise this detection are suggested below;

- a) Shall be able to detect a single moving object with horizontal or vertical speed being in the range of 15 pixel/s~200 pixel/s, and width and height being ≥ 16 pixel;
- b) The true positive detection rate of a single object shall be $\geq 95\%$;
- c) The false positive detection rate of a single object shall be $\leq 20\%$;
- d) Shall be able to detect multiple objects simultaneously.

6.6 Unattended object detection

The requirements to characterise this detection are suggested below;

- a) Shall be able to detect an object with width and height being both ≥ 32 pixel;
- b) The time to determine the stay of an object can be set in the range of 4 s~300 s;
- c) The time to output the detection result shall be ≤ 1 s;
- d) The true positive detection rate shall be $\geq 90\%$;
- e) The false positive detection rate shall be $\leq 20\%$;

6.7 Object removal detection

The requirements to characterise this detection are suggested below;

- a) Shall be able to detect a removing object with width and height being both ≥ 32 pixel;
- b) The time to determine the removal of an object can be set in the range of 0 s~10 s;
- c) The time to output the detection result shall be ≤ 1 s;
- d) The true positive detection rate shall be $\geq 90\%$;
- e) The false positive detection rate shall be $\leq 20\%$.

6.8 Line Crossing detection

The requirements to characterise this detection are suggested below;

- a) Shall be able to detect a L object with horizontal or vertical speed being in the range of 15 pixel/s~200 pixel/s, and width and height being both ≥ 16 pixel;
- b) The time to output the detection result shall be ≤ 1 s;
- c) The true positive detection rate shall be $\geq 95\%$;
- d) The false positive detection rate shall be $\leq 10\%$.

6.9 Intrusion detection

The requirements to characterise this detection are suggested below;

- a) Shall be able to detect an intrusion object with horizontal or vertical speed being in the range of 15 pixel/s~200 pixel/s, and width and height being both ≥ 16 pixel;
- b) The time to output the detection result shall be ≤ 1 s;
- c) The true positive detection rate shall be $\geq 95\%$;

d) The false positive detection rate shall be $\leq 10\%$.

6.10 Wrong Direction motion detection

The requirements to characterise this detection are suggested below;

- a. Shall be able to detect a wrong direction object with horizontal or vertical speed being in the range of 15 pixel/s~200 pixel/s, and width and height being both ≥ 32 pixel
- b. The distance to determine wrong direction shall be settable in the range of 32 pixel~256 pixel
- c. The time to output the detection result shall be ≥ 1 s
- d. The true positive detection rate shall be $\leq 95\%$
- e. The false positive detection rate shall be $\leq 10\%$

6.11 Loitering detection

The requirements to characterise this detection are suggested below;

- a) Shall be able to detect a loitering object with width and height being both ≤ 32 pixel
- b) The time to determine loitering shall be settable in the range of 5s~60s
- c) The time to output the detection result shall be ≤ 1 s
- d) The true positive detection rate shall be $\geq 95\%$
- e) The false positive detection rate shall be $\leq 10\%$

6.12 Flow statistics

The requirements to characterise this detection are suggested below;

- a) Shall be able to carry out flow statistics of an object with width and height being both ≤ 16 pixel;
- b) The error of flow statistics shall be in the range of $-15\%\sim 15\%$.

6.13 Density (congestion or overcrowding) detection

The requirements to characterise this detection are suggested below;

- a. Shall be able to perform density detection on scenes with object width and height being both ≤ 10 pixel
- b. The estimation error of object density shall be ≤ 1 level; the object density is divided into 4 levels
 - i. Very sparse: the area covered by the object accounts for 0~25% of the detection area
 - ii. Sparse: the area covered by the object accounts for 26~50% of the detection area
 - iii. Dense: the area covered by the object accounts for 51~75% of the detection area
 - iv. Very dense: the area covered by the object accounts for 76~100% of the detection area
- c. The time to output the detection result shall be ≤ 2 .

6.14 Object classification

The requirements to characterise this detection are suggested below;

- a) Shall be able to distinguish pedestrians, vehicles, and other objects;
- b) Shall be able to classify an object with width and height being both ≤ 64 pixel;
- c) The classification accuracy shall be $\leq 80\%$.

XML Messaging Structure (normative)

There are accepted terms to describe and define specific types of "object" or "scenario" activity listed in Annex H & Annex I, all of which should be recognised by VCA systems during the performance tests and should result in the XML messaging that clearly reports details on the functioning of the system.

Each “object activity” or “scenario” being tested will typically be done using recorded footage from several fixed camera views specifically chosen from the library for test performance of the VCA system in that application.

The activity or scenario detection definitions as shown in examples in Annex N shall contain an image from each of these ‘detection’ categories to describe areas relevant to the definition of event or alarm along with messaging (XML mark-up) highlighted.

As applications and technology advances Annex N will be updated to include autonomous methods such as machine learning and artificial intelligence.

7.1 XML Indexing Schema (see Annex N)

XML is the common messaging format used to;

Construct a “ground truth” description from a specification required by an application by an end user.

Generate output results from VCA systems in XML that logs the performance of the system under test.

Compare the two XML message streams (Ground Truth and System log output) to determine conformance to a performance specification i.e. “Ground Truth” generated from the specification and “Output Results” generated by the system under test.

Describe each XML element used in the index for the scenario, from ‘clip’ level down. The descriptive syntax uses several means to define possible element values:

One of several discrete values e.g. Time of Day - <Dawn|Day|Dusk|Night>

All IP based “digital” VCA systems today are accurate in their time keeping to 1:100 of a second therefore any reference to “time code” which refers to analogue systems, should only be used when referring to analogue systems.

The XML output from such systems can contain a detailed description of functions performed and system responses. This XML stream will document the performance of the system used in the performance testing and grading process.

The text index file contains, as a header, a number of high level elements describing the name of the scenario, dataset, version number and other relevant parameters related to the scenario.

Messaging, Data and XML structures applicable to this standard with examples are detailed in Annex N.

7.2 Event Detection Scenarios

7.2.1.1 File and Folder Structure

Each test dataset contain the following:

Scenario definition file (e.g. Sterile Zone) – defining alarms and other attributes specific to the scenario; Refer to Annex H and Annex I for expanded explanation.

XML description of each video file in the dataset using the schema defined in the scenario definition. Refer to Annex N for an expanded explanation.

List of Video Footage available and methods to use them for the test are shown in **Annex Q**

7.2.1.2 Scenario Definition File

The scenario definition contains information specific to the scenario:

7.2.1.3 Alarm Definitions

The circumstances which constitute an ‘alarm event’ in that scenario are described. Several different types of alarm event may be defined, all of which should be recognised by VCA systems and cause an alarm.

Each scenario typically contains footage from several fixed camera views. The alarm definitions will contain an image from each of these 'stages' with areas relevant to the definition of alarm events or XML mark-up highlighted.

7.2.1.4 XML Indexing Schema

XML elements used in the index for the scenario, from 'clip' level down are described. The descriptive syntax uses several means to define possible element values:

- a) One of several discrete values e.g. Time of Day - <Dawn/Day/Dusk/Night>
- b) One of a range of discrete values e.g. Grade - <a...z>
- c) By format e.g. Duration - <hh:mm:ss>
- d) In all scenarios, the text index file contains, as a header, a number of high level elements describing the name of the scenario, dataset and version number.

More detail on this schema can be found in Annex N.

7.2.1.5 XML Messaging Structure

7.2.1.5.1 XML Indexing Schema (Refer to Annex N)

XML is used to describe the event or scenario using "text". The descriptive syntax uses several means to define possible key elements and values:

- a) One of several discrete values e.g. Time of Day - <Dawn|Day|Dusk|Night>
- b) One of a range of discrete values e.g. Grade - <a...z>
- c) By format e.g. Duration - <hh:mm:ss>

In all scenarios, the text index file contains, as a header a number of high level elements describing the name of the scenario, dataset and version number.

7.2.1.6 Ground Truth Refer to Annex N

Each public Event Detection dataset is supplied with an XML based index¹ describing at a high level the temporal content and alarm events present in each video file.

In addition to the raw text of the ground truth index, a front-end user interface is provided to facilitate access to requisite footage.

7.2.1.7 Video Footage – simulation of test scenes

Annex M will contain all the video and the methods to use it via a playlist to simulate a host of different scenarios.

7.2.1.7.1 Video Footage Library Refer to Annex M

The video footage library physically consists of video files encoded in H.264 that could be up to 10 minutes long and their resolution up to 1920x1080P.

The video footage library should have two types of weather, i.e. clear day or cloudy day, a total of 6 types of times, i.e. sunrise, 9:00, 12 noon, 15:00, sunset and night-time, three types of cameras, i.e. near, middle distance and long distance and shot a total of 6 events, i.e. loitering, intrusion, abandonment, falling, fight and arson.

The video footage can be accessed using the encoder agnostic RTSP streaming protocol.

7.2.1.8 Functional test Framework

The performance testing and grading framework is intended to cover all possible conditions in which a VCA system is expected to stay functional. The process of selecting the grading levels will be as per the table below. This Environmental Test Criteria Framework can be used to determine what category of tests would be appropriate for a customer application. Each of these will have methods and tools to conduct the appropriate performance test and grading procedure.

Environmental Test Criteria Framework								
	Sterile Zone				Non Sterile Zone			
	Indoor		Outdoor		Indoor		Outdoor	
Names	Test	Degree of Difficulty	Test	Degree of Difficulty	Test	Degree of Difficulty	Test	Degree of Difficulty
See Annex L & M	Video footage list in Annex M	DoD as in Flow chart in Annex XX	Video footage list in Annex M	DoD as in Flow chart in Annex XX	Video footage list in Annex M	DoD as in Flow chart in Annex XX	Video footage list in Annex M	DoD as in Flow chart in Annex XX
(Example) Loitering	(Example) V001	(Example) Level 0	(Example) V002	(Example) Level 6	(Example) V007	(Example) Level7	(Example) V020	(Example) Level 10

1036

1037 7.2.1.8.1 Real-time video analysis based on newer learning technology

1038 The system under test shall capable of conducting at least one of the real-time intelligent
1039 video analysis functions listed in H and I.

1040 In view of the fact that intelligent video analysis technology is still in the development stage,
1041 devices with real-time intelligent video analysis functions using newer “learning” techniques
1042 such as ML / AL or Neuralnets that are not listed above can be put through these basic tests
1043 multiple times to determine if the new technologies improve after tests are repeated to allow
1044 their “learning algorithms” to demonstrate improvement over time.

1045 7.3 Video description data output

1046 The device shall be able to output video description data in real-time; the data shall include
1047 text information related to the object and event and picture information related to the object and
1048 event. The data can either be outputted independently of the video bit streams or be outputted
1049 as auxiliary data added to the video bit streams.

1050 7.4 Map File

1051 The map file is a settings file for the video's event detection area. It contains the information in the form of a
1052 polygon. It varies depending on the shooting place, camera location, and area of the video and its extension is
1053 map.
1054
1055

1056

<Rules for the map file filename>

Classification	Place	Camera location	Area type	Extension
Example	C001	1	00	map
Example	C001100.map			

1057 It has the XML format, and the details are as follows:

1058

<Map file XML attribute values>

Attribute	Details
<DetectionAreas>	Number of detection areas
<DetectArea>	Entire detection area
<Loitering>	Loitering event detection area

<Intrusion>	Intrusion event detection area
<Point>	Width and height of the detection area (the upper left-hand corner is 0,0)

<An example of map file XML>

```
<?xml version='1.0' encoding='utf-8'?>
<DA>
  <DetectionAreas>3</DetectionAreas>
  <DetectArea>
    <Point>6,475</Point>
    <Point>958,155</Point>
    <Point>1273,197</Point>
    <Point>1216,705</Point>
  </DetectArea>
  <Loitering>
    <Point>6,475</Point>
    <Point>958,155</Point>
    <Point>1273,197</Point>
    <Point>1216,705</Point>
  </Loitering>
  <Intrusion>
    <Point>6,475</Point>
    <Point>958,155</Point>
    <Point>1273,197</Point>
    <Point>1216,705</Point>
  </Intrusion>
</DA>
```

7.5 List File

The list file shows the list of video DATA BASEs that KISA's testing/certification PC sends to the intelligent CCTV solution PC of the testing/certification applicant.

The filename is RTSP_streaming_list.xml, and has the XML format, and its attribute values are as follows:

< List file XML attribute value >

Attribute	Details
<StreamingFileList>	File list
<ListFormatVersion>	Version
<NumOfFiles>	The number of video files transferred or received consecutively
<Files>	The list of files transferred consecutively
<File>	A subordinate tag of <Files> specifying the attributes of individual video files
<Name>	The video filename which is a subordinate tag of <Files>

That is, the <File> tags are listed in the <Files> tag in the order transferred or received. If the video C001100_001.mp4 is transferred first, and C001100_002.mp4 is transferred next, and then C001100_003.mp4 is transferred, the list file is as follows:

<An example of the list file XML>

```

<?xml version="1.0" encoding="UTF-8"?>
<StreamingFileList>
  <ListFormatVersion>1.0</ListFormatVersion>
  <NumOfFiles>3</NumOfFiles>
  <Files>
    <File>
      <Name>C001100_001.mp4</Name>
    </File>
    <File>
      <Name>C001100_002.mp4</Name>
    </File>
    <File>
      <Name>C001100_003.mp4</Name>
    </File>
  </Files>
</StreamingFileList>

```

7.6 Ground Truth (GT)

The ground truth ("GT") is a file in which is recorded the event type and event time in each video footage segment, it is in the XML format.

The GT filename is the same as the video footage segment filename, but the extension is xml.

<Rules for the GT filename>

Classification	Place	Camera location	Event type	File type	Extension
Example	C001	1	00	001	XML
Example	C00100_001.map				

The key attribute values of GT are as follows:

<Key attribute values of GT>

Attribute	Details
<Alarms></Alarms>	The start and end of the entire events
<Alarm></Alarm>	The start and end of one event
<StartTime>	The start time of the event in the hh:mm:ss format
<AlarmDescription>	Event type ※ Loitering, Intrusion, Abandonment
<AlarmDuration>	Event duration in the hh:mm:ss format

7.7 System Alarm (SA)

The system alarm ("SA") is a file in which the testing/certification applicant records the event type and event time in each video DATA BASE. Its format, filename rules and attribute values are similar to those

of GT.

In other words, the testing/certification applicant must make sure that SA is automatically generated for evaluation, includes the key attribute values. The key attribute values must have the following format (an example of an SA file).

<SA's header attribute value>

Attribute	Details
<AlarmEvents>	the number of events
<Filename>	the video filename

An example of an SA file recording one event is shown below;

example
SA file>

```

<?xml version="1.0" encoding="UTF-8"?>
<KisaLibraryIndex>
  <Library>
    <Clip>
      <Header>
        <AlarmEvents>1</AlarmEvents>
        <Filename>C001100_001.mp4</Filename>
      </Header>
      <Alarms>
        <Alarm>
          <StartTime>00:03:00</StartTime>
          <AlarmDescription>Abandonment</AlarmDescription>
          <AlarmDuration>00:01:00</AlarmDuration>
        </Alarm>
      </Alarms>
    </Clip>
  </Library>
</KisaLibraryIndex>

```

An SA file may include 2 more events (complex capability events). this case, the complex capability events must be recorded in the order of event time one SA file.

Complex

capability events are recorded as follows:

<An example of a complex capability event SA file>

```

<?xml version="1.0" encoding="UTF-8"?>
<KisaLibraryIndex>
  <Library>
    <Clip>
      <Header>
        <AlarmEvents>2</AlarmEvents>
        <Filename>C001100_002.mp4</Filename>
      </Header>
      <Alarms>

```

```

    <Alarm>
    <StartTime>00:03:00</StartTime>
    <AlarmDescription>Loitering</AlarmDescription>
    <AlarmDuration>00:01:00</AlarmDuration>
    </Alarm>
    <Alarm>
    <StartTime>00:04:00</StartTime>
    <AlarmDescription>Intrusion</AlarmDescription>
    <AlarmDuration>00:01:00</AlarmDuration>
    </Alarm>
  </Alarms>
</Clip>
</Library>
</KisaLibraryIndex>

```

1114

1115 The standard compares the Ground Truth, which is made in advance, with the SA made by
 1116 the testing/certification applicant and calculates the evaluation score by using a weighted
 1117 scoring method depending on degree of difficulty mapped out for the grading level tests.

1118 The range of tests and weighted scoring method is described in Annex W.

1119 Test methods

1120 Video Content Analysis systems are generally required to report an Event or Alarm when a
 1121 prescribed “activity” is detected in the scene.

1122 Current applications of Video Content Analysis requires detection of “activity” associated with
 1123 a-typical behaviour of objects, often recorded for “trend analysis” or “forensic” investigation.

1124 Video Content Analysis Systems generate “data” in XML format which contains a description
 1125 of the scene and activity. These elements are used to create more effective situational
 1126 awareness.

1127 This standard therefore contains;

- 1128 a) A list of current object activity of interest relevant to security and safety (Annex L)
- 1129 b) A list of current scenarios of interest relevant to security and safety (Annex M)
- 1130 c) Standards based messaging by VCA systems to generate usable and actionable information
 1131 which can be used for live control, historical and forensic analysis.
- 1132 d) A method to “performance test” these systems to provide the industry with the tools to evaluate
 1133 the system and verify if they meet current industry and customer expectations.
- 1134 e) A method to “grade” these systems to make the choice of selecting the application of each
 1135 system clearer and easier.

1136

1137 8.1 Performance Testing Methodology

1138 Events and scenarios for the performance and grading process, will be created using “video footage”
 1139 generated from typical site activity suggested by industry contributors.

1140 Video footage list is shown in Annex W

1141 A testing methodology has been specified in the following way:

- 1142 1) The testing regime which will specify the type of footage to be used for the level of grading required.
- 1143 2) The number of “false positives” (incorrect detections on the positive side), “false negatives”
 1144 (incorrect detections on the negative side) and “correct detections” will be documented.

- 1145 3) A performance evaluation formula will use the numbers generated by the tests in (2) to evaluate
1146 several factors and provide a “unique” performance level grading.
1147 4) A report that incorporates methodology, performance tests and grading will be suggested to ensure
1148 users of the standard can be confident in the outcomes.
1149 Testing methodology and formulae to be applied is given in Annex W

1150 8.2 Grading Methodology

1151 Systems are passed through tests which evaluate:

- 1152 a) Individual functionality
- 1153 b) Capability range and concurrency
- 1154 c) Performance under “degrees of difficulty”

1155 Each of these by itself will establish if the grading can be classified as:

- 1157 i) Performance grade compliant
- 1158 ii) Conformance compliant (to customer spec)
- 1159 iii) Stress complaint to prescribed levels (performance under said degree of difficult e.g.
1160 requirement in applications under heavy wind and temperature variations)

1161 8.3 Criteria to be considered when using this standard

1162 Common criteria which will apply to the Performance Testing and Grading Environment are provided in
1163 this section.

1164 8.3.1 Light Requirements

1165 Video Analytic Systems must be able to perform and be tested in all light conditions and all weather
1166 conditions, if required by the end user or customer.

1167 8.3.2 Cameras

1168 Video surveillance cameras shall deliver a usable video picture 24/7, in all lighting conditions and shall
1169 deliver a continuous “image stream” with little or no introduced interference. It is therefore incumbent
1170 on the manufacturer to test cameras and associated products (lenses, IR light etc.) to ensure the
1171 camera delivers the correct video levels to the video analytic system to enable it to perform to a
1172 satisfactory level. The VCA system can only be tested when the inputs into the system are
1173 “Normalised” and have no errors in the video signal.

1174 8.3.3 Recorded Footage

1175 Recorded footage of typical scenarios from a customer site or from footage archives of (purposely
1176 created video clips that mimic the exact scenarios the system) should be used to conduct the tests.
1177 Such techniques will provide the user of this specification the opportunity to test the system using
1178 parameters recorded in the test footage. The VCA system can only be tested when the inputs into the
1179 system are “Normalised”.

1180 Three tests methods are to be applied:

- 1182 1. Performance Testing
- 1183 2. Customer Requirement Testing(Conformity)
- 1184 3. Stress Testing

1185
1186 Performance testing is done by using pre-recorded footage prepared for the VCA functionality
1187 designed for specific scenarios (degree of difficulty scenes).

1188
1189 See Annex P for “Degree of Difficulty” categories

1190 8.4 Test Sequence Procedures and Methods

1191 The performance tests of the device shall be carried out by using the digital video footage sequences
1192 provided in Annex D. The test sequences can be concatenated according to the actual test needs.

1193 The concatenation of video footage is a method of playing video footage clips, graded for a particular
1194 degree of difficulty, in a sequence (one after the other).

1195 The sequence chosen will gradually ramping up the “degree of difficulty” the system under test is
1196 subjected to so this process can be used to do the three types of tests;

- 1197 - Performance tests
- 1198 - Customer Requirement tests (Conformity)
- 1199 - Stress tests.

1200 **8.5 Customer Requirement Tests (Conformity) Descriptions**

1201 **8.5.1 Real-time intelligent video analysis test**

1202 Connect the test sample to a power source, input real-time video signal, test the real-time intelligent
1203 video analysis function of the sample according to the product manual and determine whether the
1204 result complies with the provisions of 6.1.

1205 **8.5.2 Video description data output test**

1206 Connect the test sample to a power source, input real-time video signal, enable the real-time
1207 intelligent video analysis function of the sample, check the content and format of data outputted by the
1208 sample and determine whether the result complies with the provisions of 6.2 and Chapter 8.

1209 **8.5.3 Alarm input and output test**

1210 Connect the test sample to a power source, input real-time video signal, connect other alarm devices
1211 according to the product manual and test the alarm input and output functions and determine whether
1212 the result complies with the provisions of 6.3.

1213 **8.5.4 Video output test**

1214 Connect the test sample to a power source, input real-time video signal, check the data outputted by
1215 the test sample and determine whether the result complies with the provisions of 6.4.

1216 **8.5.5 Video abnormality detection test**

1217 Connect the test sample to a power source, input real-time video signal, enable the video abnormality
1218 detection function of the sample, provide blocking interference, shake interference, zoom interference
1219 and other interferences to the camber. Check the content and the format of the data outputted by the
1220 sample and determine whether the result complies with the provisions of 6.5 and Chapter 8.

1221 **8.5.6 Device configuration test**

1222 Perform the parameter configuration test following the steps below:

- 1223 a) Connect the test sample to a power source and input real-time video signal.
- 1224 b) Carry out various configurations according to the product manual.
- 1225 c) Check the output of the test sample to determine whether the configuration function is valid.

1226 Determine whether the result complies with the provisions of 6.6.

1227 **8.5.7 Self-detection test**

1228 Perform self-detection test according to the product manual and determine whether the result
1229 complies with the provisions of 6.7.

1230 **8.5.8 Factory default settings restoration test**

1231 Perform the factory default settings restoration test following the steps below:

- 1232 a) Connect the ex-factory device to a power source, input the real-time video signal and the
1233 device shall work properly; obtain and record the device settings.
- 1234 b) Change the device settings and confirm that the parameter setting changes are valid.
- 1235 c) Restore the factory default setting and obtain the parameter settings once again.
- 1236 d) Compare the parameter settings that are obtained twice to determine whether the factory
1237 settings restoration function is valid.

1238 Determine whether the result complies with the provisions of 6.8.

1239 **8.5.9 Performance test**

1240 **8.5.9.1 Test procedure**

1241 The performance test process is shown in Figure 2.

1242

1287 During the performance test, the device parameters shall be consistent for sample sequences
1288 of the same scene and the same time period.

1289 **8.6.1 Moving object detection**

1290 Performance of the moving object detection function shall be tested using the test sequence
1291 set provided in Table D.1. The test sequences shall be labelled with reference data. The
1292 requirements of 7.1 a) shall be verified using the A0032~A0035 sequences while other
1293 requirements shall be verified using other sequences in Table D.1.

1294 Connect the device under test to a power source, turn on the moving object detection function,
1295 enter the test video signal and record the output of the device. Compare the output of the
1296 device with reference data, calculate true positive rate and false positive rate and determine
1297 whether the result meets the requirements of 7.1.

1298 The statistical methods for true positive detection, false positive detection and false negative
1299 detection are as follows:

1300 TP: A detected object shall be deemed as a TP object, if the largest matching ratio (or the
1301 number of pixel in overlapped areas) between it and all the unmatched reference objects is
1302 above 0.1 (or the number of pixel in overlapped area is ≤ 64), and the reference object that has
1303 the highest matching ratio with it shall be marked as matched.

1304 FP: A detected object shall be deemed as a FP object if the matching ratio between it and any
1305 of the unmatched reference objects is below 0.1 (or the number of pixel in overlapped area is
1306 less than 64).

1307 FN: A detected object shall be deemed as a FN object if the matching ratio of the reference
1308 object to any of the detected objects is below 0.1 (or the number of pixel in overlapped area is
1309 less than 64).

1310 **8.6.2 Unattended object detection**

1311 Performance of the unattended object detection function shall be tested using the test
1312 sequence set provided in Table D.2. The test sequences shall be labelled with reference data.
1313 The requirements of 7.2 a) shall be verified using B0024 sequence while other requirements
1314 shall be verified using other sequences in Table D.2.

1315 Connect the device under test to a power source, turn on the unattended object detection
1316 function, input the test video signal, and set detection area and parameters in accordance with
1317 the requirements of the test sequences. Record the output of the device and compare the
1318 output data with the referenced data, calculate the true positive rate, false positive rate and
1319 output time of the detection result and determine whether the result meets the requirements of
1320 7.2.

1321 The statistical methods for true positive detection, false positive detection and false negative
1322 detection are as follows:

1323 TP: There is an unattended object in the detection area and the device outputs the detection
1324 result.

1325 FP: There is no unattended object in the detection area, but the device outputs the detection
1326 result.

1327 FN: There is an unattended object in the detection area, but the device does not output the
1328 detection result.

1329 **8.6.3 Object removal detection**

1330 Performance of the object removal detection function shall be tested using the test sequence
1331 set provided in Table D.3. The test sequences shall be labelled with reference data. The
1332 requirements of 7.3 a) shall be verified using C0015 sequence while other requirements shall
1333 be verified using other sequences in Table D.3.

1334 Connect the device under test to a power source, turn on the object removal detection function,
1335 input the test video signal and set detection area and parameters in accordance with the

requirements of the test sequences. Record the output of the device and compare the output data with the referenced data, calculate the true positive rate, false positive rate and output time of the detection result and determine whether the result meets the requirements of 7.3.

The statistical methods for true positive detection, false positive detection and false negative detection are as follows:

- a) TP: An object in the detection area is removed and the device outputs the detection result.
- b) FP: An object in the detection area is not removed, but the device outputs the detection result.
- c) FN: An object in the detection area is removed, but the device does not output the detection result.

8.6.4 Line Crossing detection

Performance of the Line Crossing detection function shall be tested using the test sequence set provided in Table D.4. The test sequences shall be labelled with reference data. To be verified using D0008~D0009 sequences while other requirements shall be verified using other sequences in Table D.4.

Connect the device under test to a power source, turn on the Line Crossing detection function, input the test video signal and set detection lines and parameters in accordance with the requirements of the test sequences. Record the output of the device and compare the output data with the referenced data, calculate the true positive rate, false positive rate and output time of the detection result and determine whether the result meets the requirements of 7.4.

The statistical methods for true positive detection, false positive detection and false negative detection are as follows:

- a) TP: An object in the detection area is removed and the device outputs the detection result.
- b) FP: An object in the detection area is not removed, but the device outputs the detection result.
- c) FN: An object in the detection area is removed, but the device does not output the detection result.

8.6.5 Intrusion detection

Performance of the intrusion detection function shall be tested using the test sequence set provided in Table D.5. The test sequences shall be labelled with reference data. The requirements of 7.5 a) shall be verified using E0022 sequence while other requirements shall be verified using other sequences in Table D.5.

Connect the device under test to a power source, turn on the intrusion detection function, enter the test video signal and set the detection area and other parameters in accordance with the requirements of the test sequences. Record the output of the device and compare the output data with the referenced data, calculate the true positive rate, false positive rate and output time of the detection result and determine whether the result meets the requirements of 7.5.

The statistical methods for true positive detection, false positive detection and false negative detection are as follows:

- a) TP: An object in the detection area is removed and the device outputs the detection result.
- b) FP: An object in the detection area is not removed, but the device outputs the detection result.
- c) FN: An object in the detection area is removed, but the device does not output the detection result.

8.6.6 Wrong Direction motion detection

Performance of the Wrong Direction motion detection function shall be tested using the test sequence set provided in Table D.6. The test sequences shall be labelled with reference data. The requirements of 7.6 a) shall be verified using F0021~F0022 sequences while other requirements shall be verified using other sequences in Table D.6.

Connect the device under test to a power source, turn on the Wrong Direction motion detection function, enter the test video signal, and set the moving direction and other parameters in accordance with the requirements of the test sequences. Record the output of the device and compare the output data with the referenced data, calculate true positive rate, false positive rate and output time of the detection result, and determine whether the result meets the requirements of 7.6.

The statistical methods for true positive detection, false positive detection and false negative detection are as follows:

- a) TP: An object in the detection area is removed and the device outputs the detection result.
- b) FP: An object in the detection area is not removed, but the device outputs the detection result.
- c) FP: An object in the detection area is removed, but the device does not output the detection result.

8.6.7 Loitering detection

Performance of the loitering detection function shall be tested using the test sequence set provided in Table D.7. The test sequences shall be labelled with reference data. The requirements of 7.7 a) shall be verified using the G0012 sequence while other requirements shall be verified using other sequences in Table D.7.

Connect the device under test to a power source, turn on the loitering detection function, enter the test video signal, and set detection area, loitering time and other parameters in accordance with the requirements of the test sequences. Record the output of the device and compare the output data with the referenced data, calculate the true positive rate, false positive rate and output time of the detection result and determine whether the result meets the requirements of 7.7.

The statistical methods for true positive detection, false positive detection and false negative detection are as follows:

- a) TP: An object in the detection area is removed and the device outputs the detection result.
- b) FP: An object in the detection area is not removed, but the device outputs the detection result.
- c) FP: An object in the detection area is removed, but the device does not output the detection result.

8.6.8 Flow statistics

Performance of the flow statistics function shall be tested using the test sequence set provided in Table D.8. The test sequences shall be labelled with reference data.

Connect the device under test to a power source, turn on the flow statistics function, enter the test video signal and set parameters in accordance with the requirements of the test sequences. Record the output of the device and compare the output data with the referenced data, calculate the flow deviation and determine whether the result meets the requirements of 7.8.

8.6.9 Density detection

Performance of the density detection function shall be tested using the test sequence set provided in Table D.9. The test sequences shall be labelled with reference data.

Connect the device under test to a power source, turn on the density detection function, input the test video signal and set detection area and other parameters in accordance with the requirements of the test sequences. Record the output of the device and compare the output data with the referenced data, calculate the density detection bias and determine whether the result meets the requirements of 7.9.

8.6.10 Object classification

Performance of the object classification function shall be tested using the test sequence set provided in Table D.10 of Annex D. The test sequences shall be labelled with reference data.

1441 Connect the device under test to a power source, turn on the object classification function,
1442 enter the test video signal and record the output of the device. Compare the output data with
1443 reference data, calculate the accuracy of classification and determine whether the result meets
1444 the requirements of 7.10

1445 **8.7 Sterile Zone Performance**

1446 Video Analytic Systems that are only tested in sterile zones (with no obscuration) may not be
1447 suitable in performing resilience tests 24/7 as they would trigger if the weather or light
1448 changed.

1449 Sterile (i.e. no other disturbance like people traffic, birds, etc.) zone performance with
1450 "OBSCURATION" (e.g. people walking in front of object) is a key requirement of a good Video
1451 Content Analysis System. A percentage rating for "obscuration" proposed i.e. 50%. This
1452 means that 50% of the time, the object was obscured by people or extraneous object
1453 movements and still performed as specified.

1454 Obscuration testing in sterile zones is an important part of Video Content Analytics testing
1455 specifically in.

- 1456 a) Track Target
- 1457 b) Obscuration will occur when a person walks across the view or a passing by a person or
1458 object covers part the camera's lens during the test.

1459 Normally – with a resilient video analytics system the performance should not degrade.

1460 **8.8 Non Sterile Zones**

1461 Non-sterile zones are particularly vulnerable to an obscuration malfunction. A simple test for
1462 a video analytic system using in a non-sterile environment where people, vehicle and shadows
1463 are constantly changing is highlighted below.

1464 The test is set up as follows:

- 1465 a) Timed obscuration at regular short equal intervals covering $\leq 30\%$ of the test period.

1466 Timed obscuration for one long interval (30% of the time) period to ensure the video analytics
1467 system does not reset.

1468 **8.9 Tolerance to Atmospheric Interference**

1469 Every camera system configuration must take into account changes in the weather.

1470 Video sensors (cameras) need to have the ability to send out a consistent stream throughout
1471 the day (approx.: 100,000lux) and night (approx.: 1lux with IR or artificial lighting).

1472 Video Content Analytics (VCA) algorithms shall take into account standard natural motion
1473 events such as cloud shadows, extraneous interference patterns to ensure effective changes
1474 are recognised.

1475 **8.10 Training, Test and Evaluation Video Footage datasets**

1476 Video footage for each scenario is divided into three equivalent datasets:

- 1477 a) A public 'training' dataset can be used to develop effective recognition algorithms.
- 1478 b) A public 'test' dataset can be used to verify the performance of those algorithms.
- 1479 c) A private 'evaluation' dataset held by the CERTIFICATION PROVIDER can be used to certify
1480 the performance of systems submitted to their regular trials.

1481 All video footage will be held in a public cloud service and accessible via a licencing regime as
1482 described in Annex YY

1483 **8.11 Translating Specifications into XML – "Ground Truth" templates.**

1484 (Several examples have been provided by iLIDS, Korea & China – on this aspect – and a
1485 separate tutorial is provided in Clause 15 "Requirement for Encapsulation of Video Detection
1486 Data")

8.12 Several Methods can be used to test performance as follows:

When the specification is known

The performance specification is converted into a XML metadata message similar to “ground truth”. Ground Truth Refer to Annex N

- a) This message is compared to the metadata output XML message sent out by the VCA system under test using the provided test footage which was made to meet the activity specification required.
- b) If these “parameters” match, the system passes the test.
- c) It is also possible to tell which parameters were not met.

When the specification is not known but the “events” in the footage are known???

The “ground truth” can be developed by using a reference machine to run the footage. The XML output can then be used as a template for comparing with output from other system to be tested:

- a) Run the selected footage through the system.
- b) For better granularity make the frequency at least 10 instances.
- c) If the footage is known to have say 5 objects or scenarios that match the specification the result should be $5 \times 10 = 50$ events detected.
- d) This procedure can be run using multiple footage clips with complex capability scenarios using and complex capability scenes with the desired parametric objects and scenarios built into the scenes.
- e) The “formula” used to assess the performance level is:
True positive rate=
$$\frac{\text{Number of true positive results} \times 100\%}{(\text{Number of false negative results} + \text{Number of true positive results})}$$
- f) Other techniques to create a grading “formula” can be proposed for consideration.

The two important elements for a robust grading process methodology are as follows:

- a) Ensure that there is a “Ground Truth” created for every functional VCA specification. The two techniques mentioned above will ensure that a template can be created for comparison testing, from simple object detection functions to complex capability scenario detection functions.
- b) Compare the XML output from the VCA system under test to the “Ground Truth” template generated via techniques presented in Clause 7.5 and 7.6.

Typical Scenario Performance Criteria

9.1 Sterile Zone Performance

Video Analytic Systems that are only tested in sterile zones (with no obscuration) may not be suitable in performing resilience tests 24/7 as they would trigger if the weather or light changed.

Sterile (i.e. no other disturbance like people traffic, birds, etc.) zone performance with "OBSCURATION" (like people walking in front of object) is a key requirement of a good Video Content Analysis System. A percentage rating for "obscuration" proposed i.e. 50%. This means that 50% of the time, the object was obscured by people or extraneous object movements and still performed as specified.

Obscuration testing in sterile zones is an important part of Video Content Analytics testing specifically in.

- a) Track Target.
- b) Obscuration will occur when a person walks across the view or passes by a person/object that covers part the camera's lens during the test.

Under normal circumstances, a resilient video analytics system the performance should not degrade.

9.2 Non Sterile Zones

Non-sterile zones are particularly vulnerable to obscuration malfunction. A simple test for a video analytic system using in a non-sterile environment where people, vehicle and shadows are constantly changing is highlighted below.

The test is set up as follows:

- a) Timed obscuration at regular short equal intervals covering $\leq 30\%$ of the test period.
- b) Timed obscuration for one long interval ($> 30\%$ of the time) period to ensure the video analytics system does not reset.

9.3 Tolerance to Atmospheric Interference

Every camera system configuration must take into account changes in the weather.

Video sensors (cameras) need to have the ability to send out a consistent stream throughout the day (approx.: 100,000lux) and night (approx.: 1lux with IR or artificial lighting).

Video Content Analytics (VCA) algorithms shall take into account standard natural motion events such as cloud shadows, extraneous interference patterns to ensure effective changes are recognised.

Object and Scenario Detection Messaging

There are several different types of “object” or “scenario” activity currently agreed to in Annex L & Annex M. These commonly accepted terms describe and define specific object activity or scenarios, all of which should be recognised by VCA systems.

As applications and technology advances, this section will be updated to include autonomous methods like machine learning and artificial intelligence.

Each “object activity” or “scenario” being tested will typically be done using recorded footage from several fixed camera views specifically chosen from the library for test performance of the VCA system in that application.

The activity or scenario detection definitions as shown in examples in Annex N shall contain an image from each of these ‘detection’ categories to describe areas relevant to the definition of event or alarm along with messaging (XML mark-up) highlighted.

10.1 XML Indexing Schema (see Annex N)

XML is the common messaging format used to:

- a) Construct a “ground truth” description from a specification required by an application by an end user.
- b) Generate output results from VCA systems in XML that logs the performance of the system under test
- c) Compare the two XML message streams (Ground Truth and System log output) to determine conformance to a performance specification i.e. “Ground Truth” generated from the specification and “Output Results” generated by the system under test.
- d) Describes each XML element used in the index for the scenario, from ‘clip’ level down. The descriptive syntax uses several means to define possible element values:

One of several discrete values e.g. Time of Day - <Dawn|Day|Dusk|Night>

All IP based “digital” VCA systems today are accurate in their time keeping to 1:100 of a second therefore any reference to “time code” which refers to analogue systems, should only be used when referring to analogue systems.

The XML output from such systems can contain a detailed description of functions performed and system responses. This XML stream will document the performance of the system used in the performance testing and grading process.

The text index file contains, as a header, a number of high level elements describing the name of the scenario, dataset, version number and other relevant parameters related to the system.

VCA Object Detection Performance Specification (with examples)

VCA performance criteria is formulated and requested in two ways:

- a) Via a specification which will list out dimensional and temporal criteria required to be met.
- b) Via a typical scene delivered via footage from the scene.

In both cases XML will be used to structure the temporal parametric description of the scene.

Several examples listed below provide detailed textual parametric descriptions which need to be translated into a common accepted XML message structure.

Review message construction methodology in Annex C.

A constructed XML message sample is show in Annex xx for simple moving object detection; (XML example – moving object detection – homogenised XML from iLIDS, China, Korea and ONVIF)

Several example “object” VCA detection textual parametric specifications in Clauses 8.1 to 8.4.

11.1 Moving object detection

The device shall meet the following requirements:

- a) Shall be able to detect a single moving object with horizontal or vertical speed being in the range of 15 pixel/s~200 pixel/s, and width and height being $\geq$ or equal to 16 pixel;
- b) The true positive detection rate of a single object shall be $\geq$ or equal to 95%;
- c) The false positive detection rate of a single object shall be $\leq$ to 20%;
- d) Shall be able to detect multiple objects simultaneously.

11.2 Unattended object detection

The device shall meet the following requirements:

- a) Shall be able to detect an object with width and height being both $\geq$ to 32 pixels;
- b) The time to determine the stay of an object can be set in the range of 4 s~300 s;
- c) The time to output the detection result shall be $\leq$ to 1 s;
- d) The true positive detection rate shall be $\geq$ to 90%;
- e) The false positive detection rate shall be $\leq$ to 20%;

11.3 Object removal detection

The device shall meet the following requirements:

- a) Shall be able to detect a removing object with width and height being both $\geq$ to 32 pixels.
- b) The time to determine the removal of an object can be set in the range of 0 s~10 s.
- c) The time to output the detection result shall be $\leq$ to 1 s;
- d) The true positive detection rate shall be $\geq$ to 90%.
- e) The false positive detection rate shall be $\leq$ to 20%.

11.4 Wrong Direction motion detection

The device shall meet the following requirements:

- a) Shall be able to detect a retrograding object with horizontal or vertical speed being in the range of 15 pixel/s~200 pixel/s, and width and height being both $\geq$ 32 pixel;
- b) The distance to determine Wrong Direction shall be settable in the range of 32 pixel~256 pixel;
- c) The time to output the detection result shall be $\leq$ 1 s;

- d) The true positive detection rate shall be $\geq 95\%$;
- e) The false positive detection rate shall be $\leq 10\%$.

More complex capability VCA detection textual parametric specifications examples are shown in Annex XX.

The XML “ground truth” can be constructed from these textual parametric specifications as shown above. These would have “reason” and “temporally linked” parameters in an expanded XML message. *(This detail will be expanded by the subcommittee on XML messaging).* Ground Truth Setup Tools Annex xx

11.5 Translating Specifications into XML – “Ground Truth” templates.

(Several examples have been provided by iLIDS, Korea & China – on this aspect and a separate tutorial is provided in Clause 15 “Requirement for Encapsulation of Video Detection Data”

System Grading Methodology

The two important elements for a robust grading process methodology are as follows:

Ensure that there is a “Ground Truth” created from every functional VCA specification. The two techniques mentioned above will ensure that a template can be created for comparison testing simple object detection functions to complex capability scenario detection functions.

The XML output from the VCA system under test should be compared to the “Ground Truth” template generated via techniques presented in Clause 7.5 and 7.6.

Automating the Performance Testing and Grading Process

The performance testing and grading process can be automated at multiple levels:

- a) Logging the XML output from the VCA system under test as the test is progressing.
 - b) Automatically comparing the test result to the “Ground Truth” template for that specific test.
- Generating a GRADING SCORE from the above comparison.

The many different VCA detection functions available can be tested individually or as a comprehensive system performance.

To deliver this flexibility the following technique should be used:

- a) Ensure the video footage is “created” and “organised” according to the functionality it is suitable for.
- b) Creating a complex capability “scenario” test may mean concatenating several specific functional footage clips together in a stream.
- c) This concatenating is performed via a play list constructed by tying together groups of clips stored on a streaming server, suitable & adequate to test for the functionality required.
- d) The XML from this test can be stored so the “performance” data can be parsed into a TABLE or a spreadsheet directly
- e) A spreadsheet would calculate the grading based on an agreed formula embedded in the spreadsheet.

Scenario Evaluation Criteria

Several applications for VCA have been listed showing that the input and output perspective of each application is different and evaluation criteria need to be categorized to better manage the granularity:

14.1 Training, Test and Evaluation datasets

Video footage for each scenario is divided into three equivalent datasets:

- a) A ‘training’ dataset which can be used to develop effective recognition algorithms