Surveillance and video analytics: factors influencing the performance

ERNCIP thematic group video analytics and surveillance

Jeroen van Rest, MSc., TNO

2015

The research leading to these results has received funding from the European Union as part of the European reference network for critical infrastructure protection project.

EUR 27852 EN

Surveillance and video analytics: factors influencing the performance

This publication is a technical report by the Joint Research Centre, the European Commission's in-house science service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

JRC science hub
https://ec.europa.eu/jrc

JRC100399

EUR 27852 EN

ISBN 978-92-79-57771-0

ISSN 1831-9424

doi:10.2788/945388

Contents

Acknowledgements

The author gratefully acknowledges the contributions, suggestions and reviews of the other members of the ERNCIP thematic group video analytics and surveillance, of the ERNCIP office and of TNO colleagues.

Abstract

Large European Union communities are working on and using surveillance systems. They are of diverse backgrounds, such as end users (e.g. law enforcement agencies and critical infrastructure owners), suppliers of surveillance products and services, academics and policymakers. One of the goals of the ERNCIP thematic group (TG) on video analytics and surveillance (VAS) is to support these communities and the interaction between them with a future proof performance evaluation methodology of video analytics, and in the future potentially also of other components of a surveillance system.

There are many factors which influence the performance of surveillance components, which is complicated as threats, vulnerabilities, assets, technologies and policies change over time. Therefore, a methodological approach is required to create common understanding among these communities regarding what these factors are and what the nature of their influence is on surveillance systems. This requires a substantial, dedicated effort during the full life cycle of surveillance systems.

Based on this common understanding, the following myths, partial truths and misconceptions can be formulated.

'There are no limits to what video analytics can do'. The consequence of this misconception is that the public has unrealistic expectations of what video analytics can do. This is also known as the 'Crime Scene Investigation' effect, because in such television series, images can be enhanced indefinitely, and during the investigation the viewpoint on the scene can still be changed. It is true that with modern technologies, such as super-resolution and various forms of sensor fusion, it is possible to generate new views on existing, combined data, but it is against the laws of nature to generate new data out of thin air.

'A camera has never caught a criminal'. Neither has a police car or a uniform. While objectively true, this statement suggests that cameras do not directly contribute to security, which is a misleading statement because the security function of cameras is both to contribute to situational awareness and to deter. They have been shown to do both, if used properly.

'Crime displacement is the Achilles heel of situational security measures'. Crime displacement is an effect on the target selection phase of a criminal, e.g. placing a camera in one area has the result of moving crime to another. Depending on the purpose of the security measures this can be a very desirable effect.

'Security only costs money'. A security department does not generate income for a business, nor should a police organisation generate net income for a state. But an investment in security measures can reduce costs and sometimes even help generate more income.

'Video analytics is not fulfilling its promise'. There can be a significant difference between expected performance and realised performance. Video analytics are being applied successfully in a growing number of scenarios. However, at the same time, there is a lot of active research, so the 'promise' is continually expanded.

'Security through obscurity is bad'. Obscurity hampers peer review, which may lead to the prolonged existence of weaknesses in the security mechanism and hinder accountability. On the other hand, obscurity poses a cost factor in the preparation of adversaries, and as such contributes to deterrence.

'Video analytics is an add-on'. Video analytics is indeed a separate capability. However, it can make or break a business case and requires significant attention to the way it is incorporated in work processes, the user interface, IT infrastructure, etc. If the use of video analytics is bolted on at the end of a system engineering process, then the chance of it contributing to the desired impact is substantially minimised. In other words, the

use of video analytics should be incorporated from the beginning, and not only at the end. Another way of putting it is 'video analytics is not plug-and-play'.

The approach of the TG VAS is to use a morphological analysis of the surveillance domain. This is a problem-solving method for high-dimensional issues with many non-quantifiable aspects that can be used to describe complex domains such as surveillance and video analytics. This report contains a description of this approach and the result of the first step of the analysis: the identification of these factors and the generic nature of their influence. Based on the overview of relevant factors for surveillance systems, it is also possible to describe the relevant factors for the subcomponent video analytics.

This approach facilitates operators of critical infrastructure and their security partners, e.g. LEAs, to describe their context, environment and threats in a specific manner. Next, they can seek interaction with security partners and industry to specify relevant capabilities that should work there. If certain elements of that solution are not yet available off-the-shelf, then further interaction with research and development (R & D) partners can be sought, for example by searching for existing video test data sets which match this challenge. If none exist, then a data set can be created based on such a specification.

This report gives a brief introduction into the topic of surveillance. Then it describes the kind of method that is required to create an understanding of the relevant factors for surveillance systems and video analytics and introduces the morphological analysis. It contains the result of applying the morphological analysis on the surveillance domain (MAS). The MAS can be used to describe the surveillance system in its context, but it is too abstract to describe differences between subcomponents. The morphological analysis is also therefore applied on the subdomain of video analytics (MAVA). They are interrelated because the video analytics process is a potential subcomponent of a surveillance system and so some of the dimensions and terms are similar. The report gives examples of how this method can be used to describe key aspects in the domains of surveillance and video analytics. The next steps are to:

- disclose and develop test data sets for relevant video analytics use cases;
- apply joint innovation early in the case of innovative use of video analytics;
- prevent and stop crises with dynamically deployed surveillance capabilities;
- develop large-scale auto-calibration in order to make video analytics robust and scalable;
- develop metadata standards to cover relevant factors influencing video analytics.

Relevant use cases for critical infrastructure protection are described in [2] using the two methods introduced in this report: the MAS and the MAVA.

1 Introduction

Large European Union communities are working on and using surveillance systems. They are of diverse backgrounds, such as end users (e.g. law enforcement agencies and critical infrastructure owners), suppliers of surveillance products and services, academics and policymakers. One of the goals of the ERNCIP thematic group on video analytics and surveillance (TG VAS or TG) is to support these communities and the interaction between them with a future proof performance evaluation methodology of video analytics, and in the future potentially also of other components of a surveillance system. For example, suppliers and research institutes want to manage the risk that solutions may not be picked up by their customers, while end users want to manage risk with new types of solutions. Policymakers need to prepare policies that help create and sustain a free, competitive and secure society in the face of new surveillance technologies and services.

This is not the first initiative of this nature. The image library for intelligent detection systems (i-LIDS) initiative in the United Kingdom was started by the Home Office and has been transferred to the Centre for the Protection of National Infrastructure (CPNI). In the VIEWER project an overview was presented of video analytics performance evaluation approaches [3], and an EU initiative was announced [4], but so far there has been no follow up. The same is the case in the Netherlands [5].

1.1 The challenge

There are many factors which influence the performance of surveillance components, which is complicated as threats, vulnerabilities, assets, technologies and policies change over time. Therefore, a methodological approach is required to create common understanding among these communities regarding what these factors are and what the nature of their influence is on surveillance systems. This requires a substantial, dedicated effort during the full life cycle of surveillance systems. As understanding of the subject matter grows, this approach should allow for corrections and extensions.

1.2 The approach

The approach of the TG VAS is to use a morphological analysis [1] of the surveillance domain. This report contains a description of this approach and the result of the first step of this analysis: the identification of these factors and the generic nature of their influence. Based on the overview of relevant factors for surveillance systems, it is possible to also describe the relevant factors for the subcomponent video analytics.

This approach facilitates operators of critical infrastructure (CI) and their security partners, e.g. LEAs, to describe their context, environment and threats in a specific manner. Next, they can seek interaction with security partners and industry to specify relevant capabilities that should work there. If certain elements of that solution are not yet available off-the-shelf, then further interaction with R & D partners can be sought, for example by searching for existing video test data sets which match this challenge. If none exist, then a data set can be created based on such a specification.

1.3 Reading guide

Shared understanding of relevant terminology is a precondition to fruitful discussion, research and design. In the domains of police, surveillance, behavioural psychology and system engineering there are several concepts which are notoriously difficult to define. At the end of this report, a section with terminology and abbreviations is included.

Chapter 2 contains a brief introduction to the topic of surveillance. Chapter 3 describes the kind of method that is required to create understanding of the relevant factors for surveillance systems and video analytics and introduces the morphological analysis. Chapter 4 contains the result of applying the morphological analysis on the surveillance

domain. The MAS can be used to describe the surveillance system in its context, but is too abstract to describe differences between subcomponents. Section 4.2 and Appendix C therefore contain the result of applying the MAVA. They are interrelated because the video analytics process is a potential subcomponent of a surveillance system and so some of the dimensions and terms are similar. Chapter 5 gives examples of how this method can be used to describe key aspects in the domains of surveillance and video analytics. Chapter 6 contains the conclusions. It also describes several myths, partial truths and misconceptions related to surveillance and video analytics. The chapter ends with an elaborate description of the recommended next steps.

2 Surveillance

This chapter gives a concise introduction to the typical components of a surveillance system. It covers a system's nature as risk management measure and its purpose and effectiveness, as well as how it can contribute to safety and security. The concept of invasiveness and monetary perspective are discussed as well as the typical ethical and legal starting points. The chapter ends with a description of the concept of compartmentalisation.

2.1 Surveillance as a risk management measure

Risk is caused by the combination of an asset that is worth protecting, a threat and vulnerability in the protection of that asset. Risk is changed or even absent if any of these three ingredients is removed or even altered, which creates security (Figure 1).

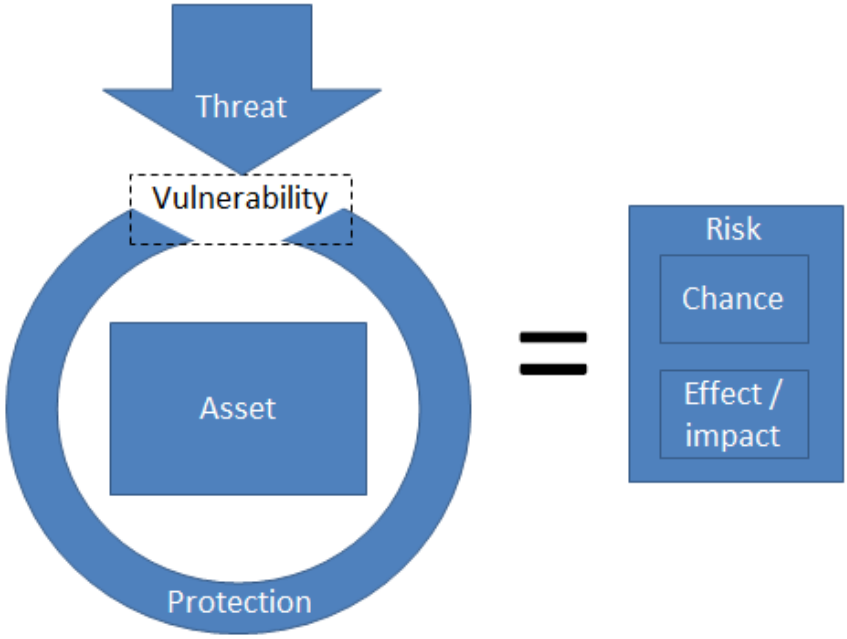

Figure 1 The combination of an asset that is worth protecting, a threat and vulnerability in the protection of that asset causes risk exposure, i.e. the chance of an impact. This figure represents a simplification of the theories of risk management

Risk management is described in ISO 31000 [6]. It defines risk as the effect of uncertainty on objectives, and this framework defines several process steps as part of risk management. Table 1 describes the potential contribution of surveillance to a subset of those steps.

Table 1 The potential contribution of surveillance to risk management processes

Risk management process	Potential contribution from surveillance
Establishing the context	• Establishing the local cultural and societal context.
Risk identification	• Identify potential sources of risk, including threats, vulnerabilities and assets to protect. • Identify potential consequences of risk, including cascading consequences.
Risk analysis	• Reveal underlying mechanisms of the causes and sources of risk. • Reveal the frequency of incidents and the extent of their effects.
Risk treatment	• Remove the risk source: detect threats and vulnerabilities before an incident, e.g. using intelligence. • Change the likelihood: deter threats. • Change the consequences: gather information for effective intervention during or after the incident, including recovery and the mitigation of cascading effects. • Share the risk: avoid liability. • Retain risk by informed decision.
Monitoring and review	• Learn. • Evaluate. • Improve risk management. • Adapt to changing circumstances.

2.2 Surveillance purpose and effectiveness

In the context of ERNCIP, surveillance is the focused, systematic and routine attention to personal detail for the purpose of protection, and more specifically for the risk treatment strategies which are supported by intelligence gathering, object security and VIP protection and crime prevention, as well as for the investigation of crime.

Surveillance can achieve this by three means: deterrence, observation and reconstruction. It can deter by increasing the chances of being caught and by revealing the modus operandi and accomplishes. This requires a minimal level of invasiveness (Section 2.4). Surveillance can detect by giving human operatives accurate and live situational awareness and/or through the use of automated processes, i.e. video analytics. Surveillance can help reconstruct an incident through the availability of footage for forensic experts, perhaps again helped by video analytics. Surveillance can also influence subjective security if surveillance resources are visible or if the consequences of surveillance can be felt. In order to determine whether surveillance technology is actually improving surveillance, the effectiveness of the latter must be expressed in terms of these higher purposes.

Simultaneously, the concrete effect of surveillance on the crime depends on the phase of the crime, e.g. the movement of crime to adjoining neighbourhoods is a proven effect of video surveillance on the target selection phase of a crime. Therefore, for surveillance technology to improve the effectiveness of a surveillance capability it must improve one or more of the specific elements of effectiveness as described in Table 2. Today's

effectiveness studies emphasise the execution phase and the risk treatment option of 'changing the consequences', such as the interruption of crime, catching the criminal in the act and criminal investigation (represented by the bold outlined cell in Table 2). However, the 'C' cells are those in which, for example, camera surveillance has been proven to be effective [7].

Table 2 Security effectiveness: the effectiveness of risk management processes expressed on criminal phases: green cells are aspects of effectiveness. Other cells are logically excluded. Preparation includes establishing the context, risk identification and risk analysis

Risk management processes / Criminal phases	Preparation	Risk treatment					Monitoring and review
		Change likelihood	Remove risk source	Change the consequences	Share risk	Retain risk by informed decision	
Developing motivation							
General target selection							
Intelligence and surveillance							
Specific target selection		C	C				
Planning and target surveillance							
Dry run							
Execution				C	C		C
Fleeing							
Enjoying the fruits of the crime							
Repentance							
Rehabilitation							

Surveillance capabilities accomplish these effects by two main means: creating situational awareness and being invasive.

2.3 Situational awareness

Surveillance systems create and maintain situational awareness by gathering and verifying relevant information on assets and on potential threats, as well as on their environment. This can include but is not limited to: the presence, amount, flow and density of groups of people; people's mood; social identity and level of intoxication; the actual location of specific people or objects and their trajectories; and even their identity

and the occurrence of deviant behaviour, including physiological properties such as sickness and excitement. The relevance of this information depends on the asset to be protected, the threats on this asset and the environment and context.

Based on the situational awareness, surveillance systems generate alerts which contribute to the threat assessment. There are several ways in which alerts can be generated from data: threshold alarm, 'bag of words' (i.e. unstructured data), concentric circles of protection (data structured per security rings), profiling and 'scenario view' [8].

Profiling differs from the other methods because it is merely a statistical assumption. This is more elaborately described in Section 6.2 of [7]. Appendix A contains taxonomy of profiling. Surveillance methods and systems are typically concerned with behaviour profiling.

2.4 Ethics, privacy and invasiveness

The ethical starting point for surveillance — or any security measure — is that there is an asset which is legitimately worthy of protection as well as a risk in relation to this asset. This risk originates from a threat and vulnerability related to the asset. Security measures thereby provide a level of freedom, or at least continuity of the asset. However, protection by definition restricts the liberties of others and often also of the asset itself, specifically their privacy. Therefore the security measures — including surveillance — should be in proportion to the importance and vulnerability of the asset and the type and level of threat against which it should be protected. Starting points of this nature are the basis for EU regulations and directives as well as for national laws.

Privacy is more than the protection of personal data. Langheinrich describes five different kinds of privacy, four of which are not concerned with personal data (Langheinrich in [9]):

- privacy of personal behaviour (media privacy),
- privacy of territory (territorial privacy),
- privacy of the person (bodily privacy),
- privacy of personal communications (interception privacy), and
- privacy of personal data (data or information privacy).

Invasiveness (intrusiveness, obtrusiveness) — in any of these kinds of privacy — is an inherent and useful aspect of surveillance, at least when applied proportionally to the context, including the threat [7]. Invasiveness can be useful because surveillance systems can prevent crime by deterrence, but this works only if the subject somehow knows of the surveillance, which is either directly because they are in its presence or personally see or experience it, or indirectly because they know or suspect that it has led to the apprehension of themselves or of someone in their social network. Research has shown that the effect of deterrence wears off unless actual follow-up occurs. Surveillance technologies can contribute to these effects, e.g. by announcing its presence and function to subjects in the scene or by being more strict (creating longer waiting lines, for example) on subjects with a criminal history, and thereby invasive.

In earlier works, five different meanings have been described regarding how the term 'invasiveness' is used in relation to surveillance [7]. These are:

- illegal surveillance, e.g. surveillance for the purpose of espionage is more invasive;
- the by-catch of other subjects, locations or moments unrelated to the threat or the asset to be protected;
- the lack of transparency, e.g. surveillance without notification is more invasive;
- the high level of detail of the data that is recorded;
- the extent to which the individual loses autonomy, i.e. has to cooperate with the surveillance.

The last two of them have also been translated into a scale of invasiveness (Table 3).

Table 3 Four- and nine-point scales of invasiveness

Invasiveness (4-point scale)		Invasiveness (9-point scale)		Definition
A	None	0	None	No surveillance
B	Light	1	Know, not seen	The subject knows he is being observed but does not see the surveillance or have to wear or do anything for this (for instance, normal camera surveillance is built into the environment).
		2	Seen	The subject sees the sensors observing, but does not have to wear or do anything.
C	Medium	3	Worn	The subject wears a device that is monitored and so must cooperate. The device requires no further action, e.g. a GPS tracking device or mobile phone.
		4	Do	The subject has to regularly do something to be monitored, such as provide biometrics in a controlled environment or present a badge to a reader.
		5	Possibly interrupt	The supervisors have the option to interrupt what the subject is doing, although this is not certain. For instance, a police officer is next to a flow of people or an access gate that is open but can close for particular subjects.
D	Strong	6	Interrupt	The subject knows that what they are doing will actually be interrupted, for example a reception desk with a waiting area at a secured building that they want to visit.
		7	Available	The subject must allow physical access to (part of) their body, as in the case of a frisk.
		8	Full transparency and cooperation	The subject allows full access to their body as well as to measure internal physiological properties.

2.5 Typical components of a surveillance system

At a minimum, a surveillance system consists of a person that combines the functions of 'sensor' and 'supervisor'. Surveillance systems can also become much more complex, depending on (perceived) threats, characteristics of the assets to protect and purpose(s) of the surveillance system. More complex surveillance systems are typically composed of components, like sensors, storage, networks, processing units, viewing stations, mobile interfaces, command and control unit(s), human supervisors and human operatives on the ground.

2.6 Compartmentalisation

Compartmentalisation is a fundamental principle of security. There are many ways to divide a space into compartments. A typical starting point is the following three-ring model.

- Ring 1 is the vital area: opponents may not reach this area.
- Ring 2 is the protected area: the purpose of this area is to allow for interventions before the opponent reaches the vital area.
- Ring 3 is the observation area: the purpose of this area is to see opponents coming.

For the purpose of this report, we consider the perimeter and the access door as separate compartments, allowing us to distinguish intrusion detection based on perimeter solutions from intruder detection based on area or volumetric solutions. We also consider the observation area (Ring 3) as a compartment. The observation area is typically a specific part of a public area and has boundaries, i.e. it is not unlimited.

A more elaborate example is given below with a secured building and a stylised picture of compartments:

a. vital area (Ring 1)
b. access door/gate
c. perimeter
d. secured area (Ring 2)
e. observation area (Ring 3)

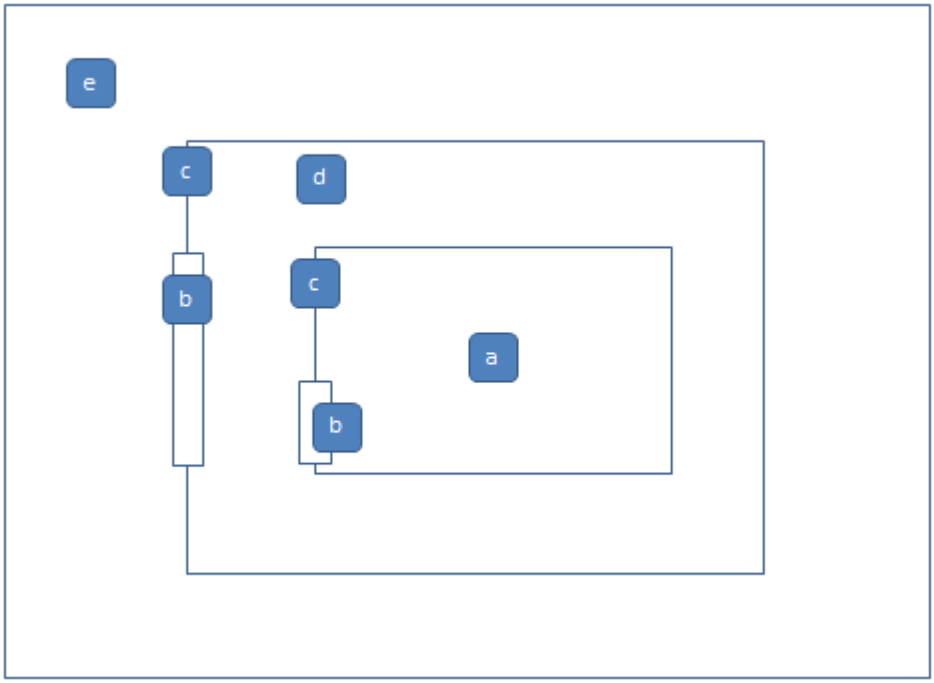

Figure 2 Compartments for object security

3 Method

The challenge of identifying relevant factors for surveillance systems and their relations has certain characteristics. These characteristics are introduced in the next paragraph.

3.1 Problem characteristics

High-dimensional

This is a high-dimensional complex problem: there are many kinds of surveillance systems and one has to take into account aspects of the system's environment and context, ethics and legal aspects, the purpose of the surveillance, the human factor both as actor and as subject, costs, technology maturity and availability, timeliness and a number of physics laws.

Quantitative and qualitative

In addition, many of these aspects cannot be quantified, such as the sociological, ethical and political factors. Even aspects which are in theory measurable may cost too much to actually determine. On the other hand, the method should be able to deal with quantified factors if they are available, such as costs, the distance between sensor and subject and the number of subjects in the scene.

Dynamic

Threats, vulnerabilities, assets, technologies and policies change over time, which means that the method must be future-proof, i.e. easy to maintain, flexible and extendable.

Varying nature of relations between factors

The relations between factors can be of varying nature. Some are based on laws of nature or other logical grounds, such as the relation between a sensor type (e.g. camera) and the modality it can observe (light). Other relations are based on empirical grounds, such as the relation between the deployment time and the mobility of a sensor: a highly mobile sensor typically has a low deployment time, but there are possible exceptions. Finally, there are also relations of a normative nature. For example, it is usually deemed too invasive to use an X-ray scanner to protect against shoplifting.

A methodology and model that can handle multidimensional data and both quantitative and qualitative factors, can deal with empirical (if available), logical (laws of nature) and normative (political and ethical) relations and is future-proof is therefore needed.

3.2 Morphological analysis

A morphological analysis (MA) is a problem-solving method for high-dimensional problems with many non-quantifiable aspects [1]. In such an analysis, the dimensions would be enumerated and several values for each dimension would be identified. Many combinations would be excluded based on logical, empirical or normative grounds. The remaining configurations are valid solutions to the problem, i.e. a configuration. A subset of dimensions (which may for some reason belong together) is a morphological box. This name comes from the 3-dimensional box that can be visualised by expressing three dimensions of a larger MA.

The selection of this method was done based on experience gained in the EU FP7 research programme TACTICS [10] [11], the Dutch national research programme 'Object security' and the Dutch TNO research programme 'Deviant behaviour' [7]. Specifically D5.1 and D5.2 of the TACTICS project describes in extensive detail a set of dimensions and values which have similarities with the set contained in this report. There are some key differences however. The scope of TACTICS is less broad than that of this TG, which leads to differences in both the dimensions and their values. For example, TACTICS is focussed on an 'urban environment', while this TG has a broader application area.

The method has limitations. It is not possible to describe structural relations between factors. For example, if both 'person' and 'vehicle' are values in the dimension 'object to be observed', then the method does not easily support the creation of the relation 'in' in the expression 'a person in a vehicle'. In a similar fashion, relations in time, such as 'before', 'during' and 'after', are not supported either. Compound threats, such as both pickpocketing and shoplifting are not supported either, nor are surveillance systems that have more than one function. A workaround can be to create separate values for frequently occurring combinations or separate dimensions for frequently occurring relations. However, these are not requirements for the purpose of understanding factors that influence surveillance and video analytics (and other subcomponents), and there are other methods (and tools) that do support them.

A drawback is the complexity to fully utilise the method. The MA requires abstract thinking and a broad perspective on the problem domain. The final step of the analysis consists of describing the relation of every combination of values, which typically results in a large matrix that may become unwieldy to handle without the proper tools. Finally, complex questions may require advanced analysis, which also requires dedicated support tools [12]. In fact, just the written expression of a configuration can already take up a lot of space.

3.2.1 Notation of a configuration

This report heavily relies on configurations of morphological analyses. A notation convention is used to keep them as brief and clear as possible:

- a configuration starts with a reference to the name of the MA, follow by a colon;
- unspecified dimensions (i.e. no values are selected) are completely omitted;
- values that are not selected are omitted;
- end of lines are omitted and the names of dimensions are underlined;
- the specification of each dimension ends with a semicolon.

Sensor	Modality	TRL	Threat direction
Camera	**Visible light**	**1**	Criminal
X-ray	Heat	**2**	Terrorist

Based on the configuration of the MAS above, the following is an example of all five rules in effect.

MAS: Sensor: camera, X-ray; modality: visible light; TRL: 1-2;

3.3 Principles for modelling surveillance systems

The way a surveillance system is modelled draws on some basic principles of security and system theories and is described in the following subsections.

3.3.1 System theory

One principle is the separation between a surveillance system and its environment. By making this boundary clear, the relation between the surveillance system and its environment can be clearly described. For example, a surveillance system has to be adapted to the environment in order to give an adequate threat assessment.

3.3.2 Levels of abstraction

The data from the (camera) sensor can be interpreted at different levels of abstraction [8] as illustrated in Figure 2. Using specific terminology for metadata on each level greatly facilitates scientific and engineering discussions.

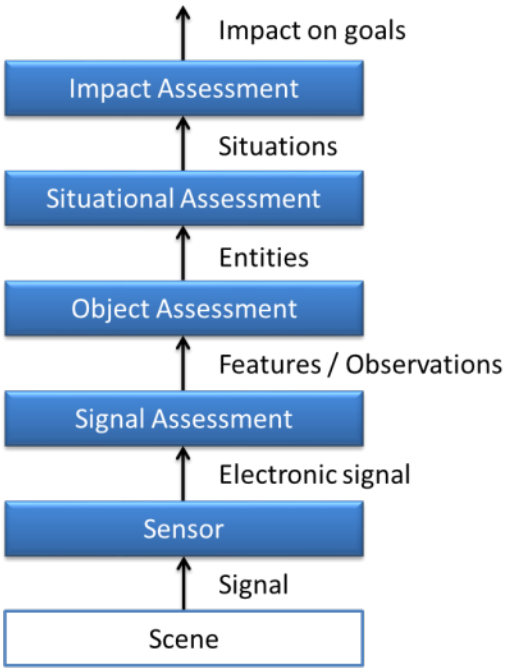

Figure 3 Multi-layered view on abstraction of information [8]

A (camera) sensor produces *signals*, which can be processed to form *observations* (synonym *features*). Observations are individual measurements. Multiple observations can be made of the same phenomenon over time, through different sensors, from different viewpoints or in different modalities. For example, the colour of an object, as observed by multiple cameras under different lighting conditions, results in different pixel colour values for the same object. Even though the colour of the real object is the same, the actual internal representation from multiple cameras is different. A surveillance system needs to be accommodated to cope with conflicting observations, possibly even benefitting from it.

3.3.3 Security as a risk management strategy

Another principle is:

1) that risk is defined as the chance that something undesirable happens; and
2) that risk is caused by the combination of:
   a) an asset to protect,
   b) vulnerability in the security system, and
   c) a threat.

$$(threat\ x\ asset\ x\ vulnerability) \implies (chance\ x\ impact)$$

Every security measure is an intervention in one or more of the three causes of risk, with the purpose of lowering the combination of chance and impact. This is a simplification of the theories of risk but is sufficient for the purpose of this paper.

3.3.4 Design basis threat

The fourth principle is the explicit design of the threat. A design basis threat is a structured approach to describe the collection of threats. This is used, for example, in the physical protection of nuclear material [13]. The specification of the threat is necessary to assess the effectiveness of a surveillance system.

The designer of a surveillance system has to consciously design the workings of a surveillance system and therefore has the freedom to ignore aspects of the environment or threat. This means that there is a difference between the design of a threat and the threat assessment as generated by a surveillance system.

3.3.5 Surveillance effectiveness

The fifth principle is that the effectiveness of a surveillance system should be specified in a combination of the incident phase and the intervention phase (Section 2.2).

3.3.6 Readiness and maturity

The sixth and final principle is that of technology readiness or system maturity. The purpose of this document is to create common understanding among scientists and end users alike. This requires at least one scale in which to express whether a technology is or should still be the subject of research or whether it is considered to be ready for use.

This is not the same as accuracy of a video analytics algorithm, for example. For some use cases, an accuracy of 20 % is sufficient, while for others a much higher level is required. The technology readiness is therefore relative to the use case.

In this report, the scale of technology readiness levels (TRL) is used, which is the most widely known scale of its kind. However, there are other related types of scales that can describe this aspect in more detail [14]:

- integration readiness level (IRL),
- system readiness level (SRL),
- concept maturity level (CML),
- innovation maturity level (IRL).

4 Results

This chapter contains the description of the actual factors as well as how they relate to each other. This is done for both surveillance systems and video analytics.

4.1 Relevant factors for surveillance systems

The categorisation of the factors is based on a model of a surveillance system in its context, as depicted in Figure 4.

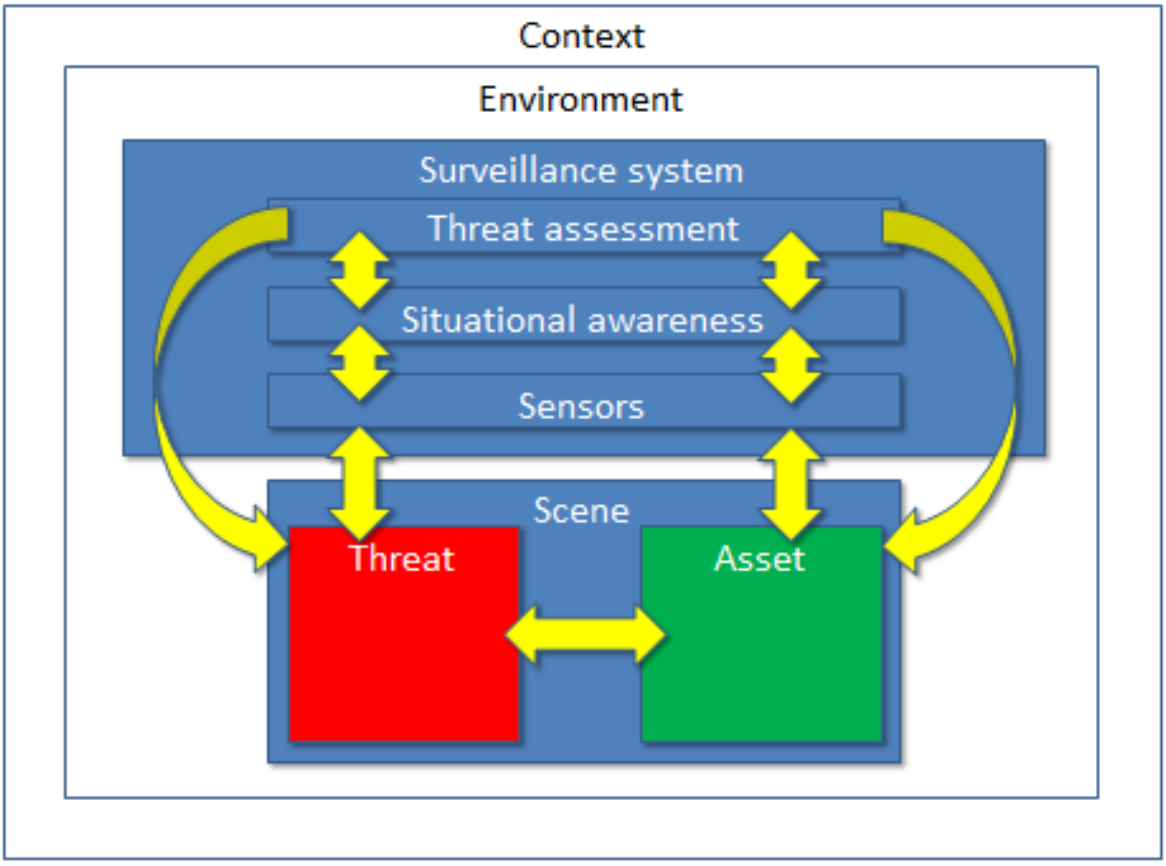

Figure 4 A surveillance system in context

The factors are grouped into categories to make them easier to understand.

- Context, environment, asset and risk:
  - context,
  - environment,
  - risk:
    - risk cause:
      - threat;
    - resulting risk.
- Surveillance system:
  - sensor,
  - situational awareness,
  - threat assessment,
  - system.

Appendix B contains a description of all dimensions and values.

Quite advanced concepts can be expressed using this method. For example, the effectiveness of a specific surveillance system might be the capability to contribute to preventing an attack in the intelligence phase.

MAS: Incident phase: before; security process: intelligence.

A rough estimate of the cost of a surveillance system could also be a dimension that can easily be related to other dimensions. For example, the mitigation of certain threats is so expensive that it cannot be done by individuals. However, the price of a concrete surveillance system can be heavily influenced by other factors which do not influence the surveillance system itself. In addition, the purpose of the TG is not to give estimate price ranges for certain surveillance systems, so this dimension is not included in this report.

4.1.1 Dependencies between surveillance factors

The factors described in Appendix B have some interdependencies. In terms of an MA, certain values of dimensions exclude values in other dimensions. For example, the platform influences the invasiveness. Table **4** shows where such exclusions can occur. This information can be used to make design decisions when designing or adapting a surveillance system and to focus research on the nature of these dependencies, for example the relation between surveillance design patterns and surveillance functions.

One situation where such a table might be helpful is when a new threat emerges. By starting with the factors that describe the threat and following the cells where dependencies are (the X symbols in Table 4), the solution space of sensible surveillance solutions can be charted.

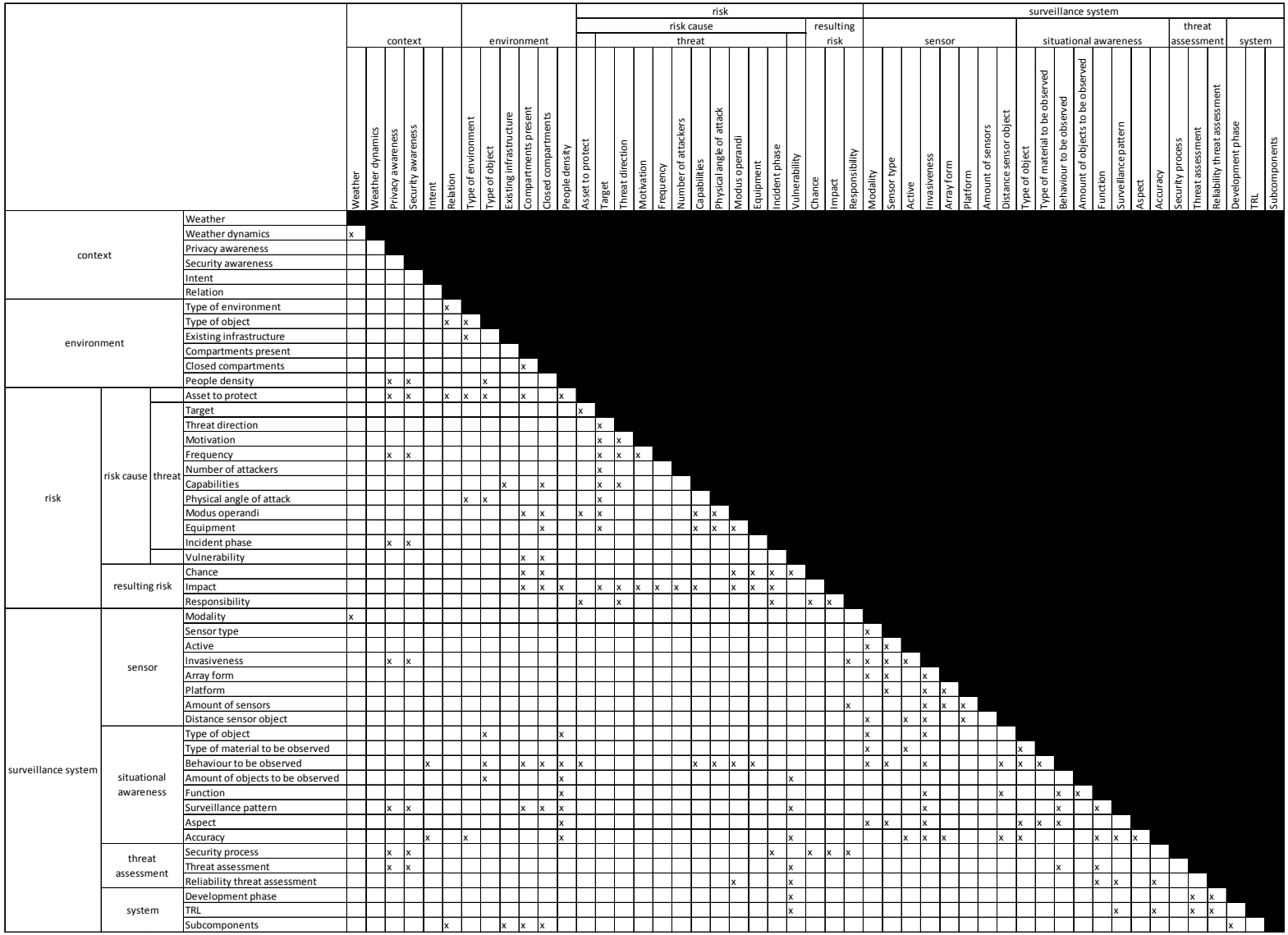

Table 4 Dependencies between surveillance factors

4.2 Relevant factors for video analytics

Video analytics is a subcomponent of a surveillance system. The scope of this section can therefore be expressed in terms of a configuration of the MAS of the previous section.

MAS: Modality: light, IR; sensor type: CCTV; active: passive, visible light (lamp), heat (infrared); invasiveness: not significant (hidden surveillance), slight (show ability to observe); development phase: configuration, use, maintenance; TRL: 7-9.

The categorisation of the factors which are relevant for video analytics is based on a slightly more specific model than that of the previous chapter, see Figure 5.

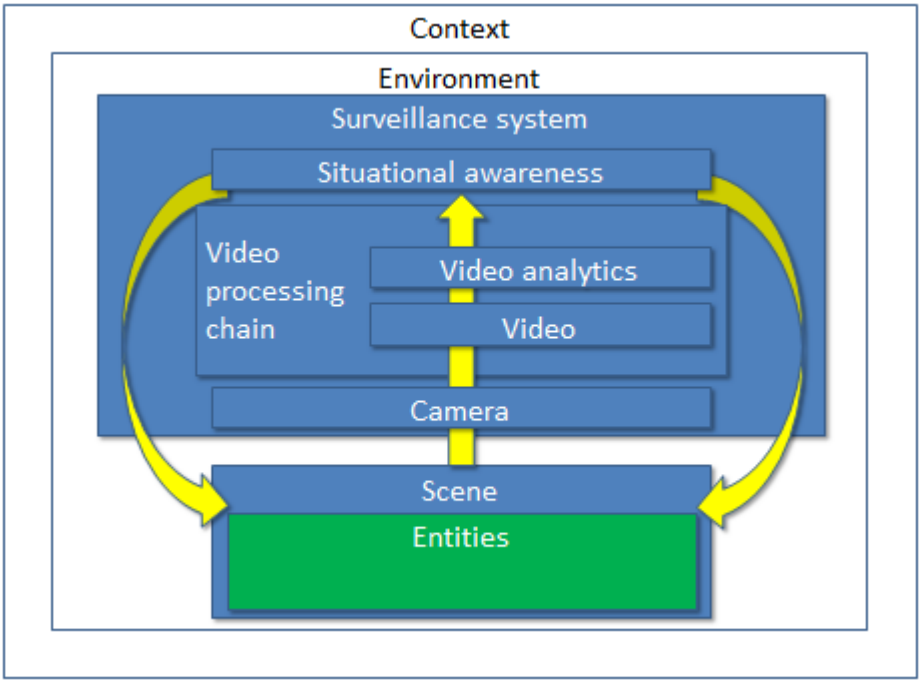

Figure 5 Video analytics in context

The relevant factors for video analytics are also grouped into categories:

— scene;
— camera;
— video processing chain:
  - video signal,
  - video analytics;
— situational awareness.

These dimensions and values are based on those of a surveillance system but may have a slightly different meaning or may be specified in more detail. Appendix C contains descriptions of all dimensions and values.

More complex functions can be expressed using this MAVA. For example, ego-motion estimation is the following.

MAVA — Platform: rotating, limited moving, free moving; object type: camera; aspect: orientation.

4.2.1 Dependencies between video analytics factors

Just like in the MAS, the factors of the MAVA described in Appendix C have some interdependencies, for example the sensor 'camera' cannot observe the modality 'X-ray'. Table 5 shows where such exclusions can occur. This information can be used to make

design decisions when designing or adapting a video analytics system and to focus research on the nature of these dependencies, for example the relation between compression and the accuracy of video analytics.

Table 5 Dependencies between relevant factors for video analytics

		Scene			Camera					Video signal		Video processing chain		Situational awareness			
		Cover	Light	Amount of objects	Distance to object	Orientation	Modality	Array form	Platform	Image improvement	Compression	Function	TRL	Object type	Aspect	Relation	Accuracy
Scene	Cover																
	Light																
	Amount of objects																
Camera	Distance to object	x	x	X													
	Orientation		x	x	x												
	Modality																
	Array form			x	x	x	x										
	Platform					x	x	x									
Video signal	Image improvement	x	x		x		x	x									
	Compression						x			x							
Video processing chain	Function	x	x	x	x	x	x	x	x	x							
	TRL							x	x	x		x					
Situational awareness	Object type							x				x	x				
	Aspect					x	x	x	x			x	x				
	Relation						x					x	x				
	Accuracy							x	x	x	x	x	x		x		

5 Instructions for use

These dimensions and values can be used to describe different key concepts for the domains of surveillance and video analytics:

— a threat or incident;
— a situation where a demand exists for new surveillance capabilities;
— the market sector that a surveillance product is designed for (and which it is not);
— what a data set represents;
— the scope of a scientific work;
— good practices.

Different communities can more easily interact with each other by using a common terminology and a basic model of a surveillance system.

The next sections contain examples of what can be described using these MA's: an actual incident, a use case for surveillance technology, a test data set and a scientific work.

5.1 Describing a threat or an incident

Describing (potential) threats and incidents that the surveillance system helps to mitigate creates clarity in the discussion between suppliers and users of surveillance systems. The next subsection contains an example.

5.1.1 Example incident: Boston bombings

The Boston Marathon bombing [15] and related shootings were a series of attacks and incidents that began on 15 April 15 2013 when two pressure-cooker bombs exploded during the Boston Marathon. Using the MAS, the bombings can be described as follows.

MAS: Weather: clouded; weather dynamics: slow changing; security awareness: medium (there was an event); intent: heterogeneous and low intensity; relation: same; type of environment: city, open outdoor; type of object: street; existing infrastructure: high; compartments present: observation area; closed compartments: none; people density: high; asset to protect: public order, the life and wellbeing of a person (e.g. the athletes and some officials), the wellbeing of a crowd, the continuation of a process (i.e. the marathon); target type: crowd; threat direction: terrorist; motivation: political; frequency: multiple; number of attackers: single; capabilities: weapons; physical angle of attack: ground; modus operandi: bombing; equipment: explosive; incident phase: during; vulnerability: intelligence, security awareness; chance: fact; impact: high; responsibility: national;

5.2 Describing use cases

The use case of a surveillance system describes the situation in which a surveillance system can be used and the threat assessment that it can generate, while remaining agnostic as to how this is done. In this way, different surveillance systems — or their subcomponents — can be tested fairly.

5.2.1 Example use case: crowd control from airborne platform

An example of a use case for a surveillance system is the surveillance of crowds from airborne platforms for crowd control. This is the assessment of relevant crowd properties for safety and security purposes. In this particular example, a use case is described where threats on crowds can be assessed from the air, using an unspecified combination of sensors and other subcomponents.

Using the MAS, it could be described as follows.

MAS: Weather: rain, clear, snowfall, clouded; weather dynamics: stable, slow changing; privacy awareness: medium (limited privacy law); security awareness: medium; intent: homogeneous and low intensity, heterogeneous and low intensity; relation: the same; type of environment: built high-rise, built low-rise, rural; type of object: street; existing infrastructure: low, medium, high; compartments present: observation area; closed compartments: none; asset to protect: public order, the life and wellbeing of a person, the wellbeing of a crowd; target: none, VIP, individual, crowd; threat direction: accident, nature, illness, activist, criminal; motivation: none, political, economic; frequency: multiple and persistent; number of attackers: none, individual, group, crowd; capabilities: none, capacity; physical angle of attack: ground; modus operandi: demonstration, molest, vandalism, theft; equipment: none, banner; incident phase: before, during, after; vulnerability: access control, security awareness, sensor coverage; chance: low, medium; impact: medium, high; responsibility: private industry, regional, public, national; active: passive; invasiveness: not significant, slight; array form: single, movement; platform: free moving; distance sensor–object: 100 m; type of object: individual, group; type of material to be observed: biological, cloth; amount of objects to be observed: 10 000; function: observe, detect; surveillance pattern: threshold alarm, bag of words; aspect: presence, behaviour, identity; security process: prevent, in the act, investigation; threat assessment: number of attackers, threatening objects or persons, modus operandi, physical angle of attack, incident phase.

5.3 Describing data sets

Data sets of surveillance videos represent a variety of real-life scenarios and as such can function as a benchmark for surveillance products, giving an early indication of the performance in real-life scenarios.

5.3.1 Example data set: i-LIDS sterile zone

The image library for intelligent detection systems (i-LIDS) is the United Kingdom government's benchmark for video analytics (VA) systems developed in partnership with CPNI [16]. There are currently five scenarios within i-LIDS, one of which considers sterile zones, where systems must detect the presence of persons in a restricted area or 'sterile zone'.

Using the MAS, it can be described as follows.

MAS: Weather: clouded, rain, clear, snowfall; weather dynamics: stable, slow changing; type of environment: rural, open outdoor; compartments present: perimeter (fence), observation area; closed compartments: perimeter, secured area (Ring 2); people density: not significant, low; asset to protect: the integrity of an object; number of attackers: single; capabilities: none, materials; physical angle of attack: ground; equipment: none, tool, camouflage; responsibility: national; modality: visible light; sensor type: camera; active: passive; invasiveness: none, slight; array form: single; platform: fixed; amount of sensors: one; distance sensor–object: 10, 100; type of object: individual; type of material to be observed: biological, cloth; behaviour to be observed: trespassing; amount of objects to be observed: one; function: detection; surveillance pattern: concentric circles of protection; aspect: presence; accuracy: 20, 50, 90, 99; threat assessment: number of attackers, threatening objects or persons, modus operandi, target, physical angle of attack, incident phase; reliability threat assessment: 90, 99; development phase: use; TRL: 8-9; subcomponents: processing unit.

5.4 Describing scientific work

Both the MAS and the MAVA can also be used to describe the topic of scientific work, which allows for quick communication of the relevance of a scientific work for non-scientists and scientists alike.

5.4.1 Example: Visual surveillance for moving vehicles

A well-known paper [17] about visual surveillance of moving vehicles from a moving platform has the following abstract.

> 'An overview is given of a vision system for locating, recognising and tracking multiple vehicles using an image sequence taken by a single camera mounted on a moving vehicle. The camera motion is estimated by matching features on the ground plane from one image to the next. Vehicle detection and hypothesis generation are performed using template correlation, and a 3D wire-frame model of the vehicle is fitted to the image. Once detected and identified, vehicles are tracked using dynamic filtering. A separate batch-mode filter obtains the 3D trajectories of nearby vehicles over an extended time. Results are shown for a motorway image sequence.'

The description based on the full article according to the MAS would be the following.

MAS: Type of environment: road; type of object: vehicle; closed compartments: none; people density: none, low; asset to protect: the life and wellbeing of a person, the integrity of an object; target: none; threat direction: accident; motivation: none; frequency: persistent; number of attackers: none; capabilities: weapons (a car); physical angle of attack: ground; modus operandi: molest; equipment: vehicle; incident phase: before; chance: medium, high, fact; impact: low, medium, high; responsibility: individual; modality: visible light; sensor type: camera; active: passive; invasiveness: not significant, slight; array form: movement; platform: limited moving; amount of sensors: one; distance sensor–object: 10 m, 100 m; Type of object: motorcycle, car, van, bus; type of material to be observed: metal; amount of objects to be observed: one; function: observe, detect; surveillance pattern: threshold alarm; aspect: presence; security process: prevent; threat assessment: threatening objects or persons; development phase: use; TRL: 4; subcomponents: sensor, processing unit;

The description of this paper using the MAVA would be the following.

MAVA: Cover: outdoor; light: none; amount of objects: one; distance to object: 10 m, 100 m; Orientation: skim; modality: visible light; array form: single; platform: limited moving; image improvement: stabilisation; compression: none; function: detection, tracking, calibration; TRL: 4; object type: vehicle, camera; aspect: presence, location, orientation; relation: spatial;

5.4.2 Example: event detection and analysis from video streams

Another example is [18] about tracking objects from an airborne platform. Its abstract is the following.

> 'We present a system which takes as input a video stream obtained from an airborne moving platform and produces an analysis of the behaviour of the moving objects in the scene. To achieve this functionality, our system relies on two modular blocks. The first one detects and tracks moving regions in the sequence. It uses a set of features at multiple scales to stabilise the image sequence, that is, to compensate for the motion of the observer, then extracts regions with residual motion and uses an attribute graph representation to infer their trajectories. The second module takes as input these trajectories, together with user-provided information in the form of a geospatial context and goal context to instantiate likely scenarios. We present details of the system together with results on a number of real video sequences and also provide a quantitative analysis of the results.'

The description based on the full article according to the MAS would be the following.

MAS: Type of environment: road, rural; type of object: long infrastructure, street; existing infrastructure: none; compartments present: access door/gate; observation area (Ring 3); closed compartments: none; people density: none, low; asset to protect: the integrity of an object; physical angle of attack: ground; equipment: vehicle; modality: visible light; sensor type: camera; active: passive; invasiveness: not significant; array form: single; platform: free moving; amount of sensors: one; distance sensor–object: 100 m, 1000 m; Type of object: car; type of material to be observed:

metal; behaviour to be observed: tailgating; amount of objects to be observed: one, two; function: observe, detect, classify; surveillance pattern: scenario view; aspect: presence, behaviour; threat assessment: threatening objects or persons, modus operandi; development phase: use; TRL: 3-4; subcomponents: sensor, processing unit;

The description of the identical paper using the MAVA would be:

MAVA — Cover: outdoor; light: none; amount of objects: one; distance to object: 100 m, 1000 m; Orientation: top down; modality: visible light; array form: single; platform: free moving; image improvement: stitching, stabilisation; function: observation, detection, tracking, calibration, shape recognition, classification, recognition; TRL: 3-4; object type: vehicle, camera, zone; aspect: presence, location, behaviour, orientation, class; relation: temporal, spatial, interaction;

6 Conclusions, discussion and next steps

This report describes the methodological approach and resulting model which captures the current relevant knowledge on the factors that influence the surveillance and video analytics. It covers the most important factors that influence the performance of surveillance and specifically those that influence the performance of video analytics and the perception thereof.

6.1 Myths, partial trusts and misconceptions

Using this knowledge, several myths surrounding surveillance and video analytics can be dispelled.

'There are no limits to what video analytics can do'. The consequence of this misconception is that the public has unrealistic expectations of what video analytics can do. This is also known as the 'Crime Scene Investigation' effect, because in such television series, images can be enhanced indefinitely, and during the investigation the viewpoint on the scene can still be changed. It is true that with modern technologies, such as super-resolution and various forms of sensor fusion, it is possible to generate new views on existing, combined data, but it is against the laws of nature to generate new data out of thin air.

'A camera has never caught a criminal'. Neither has a police car or a uniform. While objectively true, this statement suggests that cameras do not directly contribute to security, which is a misleading statement because the security function of cameras is not only to contribute to situational awareness but also to deter. They have been shown to do both, if used properly.

'Crime displacement is the Achilles heel of situational security measures.' Crime displacement is an effect on the target selection phase of a criminal. Depending on the purpose of the security measures this can be a very desirable effect.

'Security only costs money'. A security department does not generate income for a business, nor should a police organisation generate net income for a state. But an investment in security measures can reduce costs and sometimes even help generate more income.

'Video analytics is not fulfilling its promise'. There can be a significant difference between expected performance and realised performance. Video analytics are being applied successfully in a growing number of scenarios. At the same time, there is a lot of active research, so the 'promise' is continually expanded.

'Security through obscurity is bad'. Obscurity hampers peer review, which may lead to the prolonged existence of weaknesses in the security mechanism and hinder accountability. On the other hand, obscurity poses a cost factor in the preparation of adversaries, and as such contributes to deterrence.

'Video analytics is an add-on'. Video analytics is indeed a separate capability. However, it can make or break a business case and requires significant attention to the way it is incorporated in work processes, the user interface, IT infrastructure, etc. If the use of video analytics is bolted on at the end of a system engineering process, then the chance of it contributing to the desired impact is substantially minimised. In other words, the use of video analytics should be incorporated from the beginning, and not only at the end. Another way of putting it is 'video analytics is not plug-and-play'.

6.2 Discussion

This chapter discusses some consequences of choices that were made during the creation of this report.

6.2.1 Descriptions

The descriptions of dimensions in the Appendices A and B are quite descriptive, but in the current version there is no description of each value within a dimension. This may become a cause of confusion and could be part of future work.

6.2.2 Specificity

Both the dimensions and the values of the MAVA (and MAS) have currently only been described on a generic abstraction level (technology agnostic). While this contains enough detail for the description of use cases [2], this leaves out a lot of factors which become relevant when actually specifying or building a concrete surveillance system. Relevant factors which should be included in future versions are, for example:

- scene: clutter, illumination strength, illumination changes, reflections, shadows, movement in background, occlusion (both static and dynamic), speed of objects of interest, posture or orientation of object;
- sensor: field of view, camera frame rate, image resolution, perspective distortion;
- situational awareness: primitive events/actions;
- video analytics: processing platform, network bandwidth;

6.2.3 Cross-correlation matrix

The MA has a third step: to create the cross-correlation matrix, which is to determine for each value pair the statistical correlation to which they exclude each other. This exclusion can be based on empirical, normative or logical grounds. The value of this step is currently limited for the TG, so this matrix is only made on the level of dimensions for the MAVA.

6.3 Next steps

Assuming that this report contributes to generating common understanding, the question remains: what are the next steps?

6.3.1 Disclose and develop test data sets for relevant video analytics use cases

The availability of relevant test data sets is very important for all supply chains that video analytics is a part of. For law enforcement agencies they are a means to help focus the development of technologies towards actual relevant use cases. For critical infrastructure owners, test data sets help narrow the gap between expected and real value. These interests align with those of citizens, industry and scientists.

However, there are also concerns with regard to test data sets. Citizens may worry about privacy and by extension the owner of critical infrastructure. LEAs and critical infrastructure owners may worry that ongoing investigations may be hurt by releasing data sets too early and that adversaries learn security details from the test data sets, such as the positioning of cameras. Industry may worry that their investments in creating data sets are helping their competition.

Their interests and concerns are different, sometimes potentially conflicting. No one type of party has all the requirements (interest, capability and knowledge) for singlehandedly creating good in-house data sets.

While the relevance of good data sets is clear in general, it is not likely that one type of stakeholder will consistently create and maintain good, relevant data sets, with the incidental exception of some scientists (e.g. PETS). However, since technology, threats and society constantly change, this is not sustainable for scientists. When scientists and emergency services cooperate, a larger set of benefits can be aligned, creating a more sustainable solution (e.g. i-LIDS). However, without a clear chain to (monetary) business benefits, the sustainability of that solution is not high enough in terms of budget cuts

and austerity. This will not improve in the foreseeable future, so it is necessary to cooperate with more types of partners, especially those that actually earn money with VA: CI operators and/or industry. Within the ERNCIP project TG VAS, the EU has recognised this challenge and is willing to give organisational support. The vision of the TG VAS is that communities of CI end users or of industrial branch organisations develop repositories of what is for them relevant data sets because this will give them clear business continuity and monetary benefits, if only they can acquire the knowledge to do so. The TG VAS is composed of a mix of representatives of industry, scientists and emergency services. Together, this group has all the knowledge to write, for example ([1]):

(1) a clear argumentation why data sets matter in the boardroom of CI end users and industry;
(2) a manual for creating high-quality relevant data sets;
(3) a method for finding and describing relevant use cases, a step that is described in [2];
(4) a manual for maintaining and creating a repository of data sets to be used by scientists, industry and CI end users and validated by emergency services and CI end users;
(5) a procurement framework to be used by CI end users when procuring VA, making use of these data sets and repositories.

This work must be done within the context of test and evaluation frameworks such as are being developed in the ECORYS SECERCA, EU FP7 HECTOS and EU FP7 CRISP projects.

6.3.2 Joint innovation

At a certain maturity level, test data sets are no longer sufficient. Therefore, the pitfall is using them too late in the innovation process. The impact of using video analytics on a business can be substantial and a learning curve is certainly involved. If the respective parties, CI operators and LEAs, do not already have substantial experience with video analytics, then it is wise to gradually introduce video analytics in the working environment. This allows the business to adapt at their own rhythm and gives time to industry and R & D to adapt and improve the capabilities as required. This is called joint innovation.

In real working conditions, the threat that the new capability should address may not manifest itself frequently. However, representative data should obviously be generated to test, develop and adapt the capability. It may therefore be required to start using red teams during this gradual introduction.

6.3.3 Prevent and stop crises with dynamically deployed surveillance capabilities

In urban environments and around critical infrastructures there are surveillance capabilities to deal with day-to-day business, incidents and threats. However, when there is a significantly higher and more specific threat in such an environment, like a (series of) terrorist attack(s), these ordinary surveillance capabilities are no longer sufficient. Under such circumstances, LEAs have the ethical and typically also legal opportunity to use more invasive capabilities to stop that threat.

The use of a method similar to the MAVA has been validated (TRL 5) in this scenario in the EU FP7 TACTICS project. This method was used to dynamically specify, construct and deploy fit-for-purpose surveillance capabilities and, where possible, constructed from existing surveillance resources such as personnel and CCTV [10]. This is made

([1]) These are merely suggestions for now.

possible by implementing the MAVA in a decision support tool [11]. By using such a tool, it is possible to ad hoc generate surveillance capabilities in a very short space of time, without having the need to have such invasive capabilities available all the time. This is a form of privacy by design for counter-terrorism and could also be applied to other types of threats and crises.

6.3.4 Develop large-scale auto-calibration for robust and scalable video analytics

As described in Sections **Error! Reference source not found.** and 4.2, there are many relevant factors for the quality of video analytics. Not paying enough attention to these factors will lead to fragile video analytics, i.e. video analytics that will fail quickly and often in real-life situations.

If, on the one hand, the specific properties of these factors under which a particular type of analytics yields reliable performance are known and if, on the other hand, there is an efficient method to automatically determine the values of these properties in real-world circumstances, then it is possible to automatically find and perhaps even create situations with reliable performance of video analytics.

This method is called auto-calibration. Auto-calibration can help to scale up video analytics deployments while still achieving a reliable performance. This is very convenient in large-scale CCTV deployments where maintenance and moving threat patterns lead to continuous changes in the CCTV conditions. It also helps in the development of new algorithms to stay in control of the conditions under which the new type of analytics is being deployed. The third use case for auto-calibration is that of the previous section. In the case of a crisis, such as a terrorist attack, LEAs suddenly gain (legal) access to many more surveillance cameras, which are not employed under normal circumstances, whereby they are used to sustain a certain type of capability and therefore may be equipped with different types of analytics — or none at all — than are required during a crisis.

In the context of surveillance and specifically cameras, auto-calibration is a (partly) automated process to determine metadata about one or more sensors, such as a camera, from the data of the respective sensor itself. This includes its position and orientation but can also be other properties such as sensitivity, colour balance, etc. Because it is automated, it can be used to efficiently:

— deploy and maintain a large-scale network of several sensors, such as a surveillance system with cameras;
— filter sensors that are suitable for a specific capability (e.g. the tracking of persons);
— assess which capabilities are possible given a specific set of sensors.

The MAVA is as follows.

MAVA: Amount of objects: 1000; Function: calibration; object type: camera; aspect: location, orientation;

This kind of auto-calibration is generally still the topic of research. The closest functionality may be ‘tampering detection’, which is the automatic detection of whether some properties of a sensor have been changed.

It is recommended to stimulate the development into large-scale auto-calibration, especially in the light of preventing and stopping terrorist attacks and other types of crises.

Appendix F contains more generic background information about auto-calibration.

6.3.5 Develop metadata standards that cover these factors

If these (and others in Section **Error! Reference source not found.**) are indeed the relevant factors for the quality of surveillance systems and of video analytics, then they should be addressed in metadata standards [8]. This is currently not the case. Combinations of metadata standards that cover more relevant factors should be developed.

6.3.6 Other potential applications of the MAS and MAVA

The methodology introduced in this report can be used to describe incidents, security measures, projects, data sets, products, R & D outcomes, use cases and the relations between them. In [2], we apply the method to the description of contemporary use cases for video analytics.

In future works, the method could be used in research calls to describe the outcome of R & D projects, such as those of the European Commission (e.g. H2020), and to describe existing video analytics evaluation data sets.

6.3.7 Potential topics other than video analytics

Online surveillance, i.e. the surveillance of online communication networks (e.g. Facebook, email exchange, Twitter, etc.), can also be expressed using this method. However, this is more akin to data interception than to the classical observation. Just like a specialised MAVA was made to accommodate video analytics, a specialised morphological analysis could be made for online surveillance.

Other candidates for specialised MAs could be predictive behavioural profiling, command and control rooms, body cams [19] or other wearable technologies or unmanned vehicles such as unmanned aerial vehicles (UAVs).

References

[1] ERNCIP Thematic Group Video Analytics & Surveillance, "Surveillance use cases: Focus on video analytics," JRC, Ispra, 2015.

[2] FORMIT, "VIEWER - Video Intelligence Surveillance in Europe," FORMIT, 2011.

[3] A. Massimiliano and F. Fabio Bisogni, "Video Analytics: Opportunity or Spoof Story? The State of the Art of Intelligent Video Surveillance.," in *Intelligence and Security Informatics Conference (EISIC), 2011 European*, 2011.

[4] J. Van Rest, "Terminologie en Taxonomie van Video Content Analyse (Terminology and taxonomy of Video Content Analysis)," TNO, 2010.

[5] T. Ritchey, "General Morphological Analysis A general method for non-quantified modelling.," 2002.

[6] ISO, ISO 31000 Risk Management, ISO, 2009.

[7] J. Van Rest, M. Roelofs and A. Van Nunen, "Deviant behaviour - Socially accepted observation of behaviour for security - Summary," TNO, 2014.

[8] J. Van Rest, F. Grootjen, M. Grootjen, R. Wijn, O. Aarts, M. Roelofs, G. Burghouts, H. Bouma, L. Alic and W. Kraaij, "Requirements for multimedia metadata schemes in surveillance applications for security.," *Multimedia Tools and Applications 70.1,* pp. 573-598, 2014.

[9] J. Van Rest, D. Boonstra, M. Everts, M. Van Rijn and R. Van Paassen, "Designing privacy-by-design.," *Privacy Technologies and Policy. Springer Berlin Heidelberg,* pp. 55-72, 2014.

[10] TACTICS Consortium, "D3.1 TACTICS Conceptual Solution Description," 2013.

[11] TACTICS Consortium, "D5.2 Matching Capabilities to Needs," 2014.

[12] T. Ritchey, "Problem structuring using computer-aided morphological analysis.," *Journal of the Operational Research Society 57.7,* pp. 792-801, 2006.

[13] S. Isaksson and T. Ritchey, "Protection against sabotage of nuclear facilities: Using morphological analysis in revising the design basis threat." Adapted," *44th Annual Meeting of the Institute of Nuclear Materials Management,* 2003.

[14] D. Dent and B. Pettit, "Dent Associates," 2011. [Online]. Available: http://www.dentassociates.co.uk/pdf/Technology%20and%20Market%20Readiness%20Levels.pdf. [Accessed 23 December 2015].

[15] FBI, "Updates on Investigation Into Multiple Explosions in Boston," 2014.

[16] Home Office CAST, "image library for intelligent detection systems," 2010.

[17] J. M. Ferryman, S. J. Maybank and A. D. Worrall, "Visual surveillance for moving vehicles.," *International Journal of Computer Vision 37.2,* pp. 187-197, 2000.

[18] G. Medioni, I. Cohen, F. Brémond, S. Hongeng and R. Nevatia, "Event detection and analysis from video streams.," *Pattern Analysis and Machine Intelligence, IEEE Transactions on 23.8,* pp. 873-889, 2001.

[19] H. Bouma, J. Baan, F. Ter Haar, P. Eendebak, R. Den Hollander, G. Burghouts, R. Wijn, S. Van den Broek and J. Van Rest, "Video content analysis on body-worn cameras for retrospective investigation," *Proc. SPIE, vol. 9652,* 2015.

[20] J. Burgoon, R. Parrott, B. Le Poire, D. Kelley, J. Walther and D. Perry, "Maintaining and restoring privacy through communication in different types of relationships.," *Journal of Social and Personal Relationships 6.2,* pp. 131-158, 1989.

[21] S. Gutwirth, Privacy and the information age., Rowman & Littlefield, 2002.

[22] M. Langheinrich, "Privacy by design—principles of privacy-aware ubiquitous systems.," *Ubicomp 2001: Ubiquitous Computing. Springer Berlin Heidelberg,* 2001.

[23] D. Lyon, Surveillance studies: An overview., Polity, 2007.

[24] INCOSE, "A Consensus of the INCOSE Fellows," [Online]. Available: http://www.incose.org/practice/fellowsconsensus.aspx. [Accessed 8 November 2014].

[25] G. Godwin, Criminal psychology and forensic technology: A collaborative approach to effective profiling., CRC Press, 2010.

[26] Home Office, "CCTV Operational Requirements Manual 2009," 2009.

[27] R. Den Hollander, H. Bouma, J. Baan, P. Eendebak and J. Van Rest, "Automatic inference of geometric camera parameters and inter-camera topology in uncalibrated disjoint surveillance cameras," *Proc. SPIE, vol. 9652,* 2015.

List of abbreviations and definitions

Term	Definition
Agent	An agent is an autonomous entity such as a human, an animal or an automated self-controlled system (a robot). In this report it means a person in the role of victim, witness, perpetrator or supervisor.
Asset (to be protected)	The object, person, situation or process of which the continuity must be protected. This can be the life and wellbeing of a VIP, democratic order or public order in general.
Behaviour	The reaction of a cognitive agent to a stimulus, expressed in elements of its environment.
Behaviour profiling	The extrapolation of information about an agent or a group of agents, based on their behaviour.
Cognition	The ability to solve problems.
Compartment	A compartment is a conceptual subsection of a physical space. In the context of object security, it is typically enforced with security measures.
Context	The context of a surveillance system consists of the factors that influence the system and necessarily comprises the environment and the people in it. Typical examples of a surveillance context are the local culture, the level of threat and the weather conditions. Additionally, knowledge such as prior probability and known correlations between events and actions are also part of a surveillance system's context.
Effectiveness	The degree to which a desired effect is obtained.
Environment	(1) The environment of a system is the system's surroundings that could interact with the system. The typical environment of a surveillance system is the area under surveillance, including the people and the location(s) of the system components (storage, data transport, monitoring room, etc.). (2) The environment of a subject comprises the factors that directly interact with it.
Intent	The state of mind of a cognitive agent (person) that is directed towards an object or situation in their environment.
Invasiveness/ intrusiveness/ obtrusiveness	The type and degree to which the integrity of a person is breached. This has both an objective and a subjective component. However, there is no common definition of the invasiveness of a surveillance capability, which makes it difficult to answer questions such as 'how invasive is a particular surveillance capability?' or 'which is the least invasive manner of detecting a specific modus operandi?' See Section 2.4.
Object	(1) Object to be secured. (2) Object to be observed.

Term	Definition
	(3) In a surveillance system, the internal representation of an object.
Privacy	The definition of privacy is not settled. Privacy is the ability to control and limit physical, social, psychological and informational access to the self or one's group [20]. Gutwirth writes that privacy is the safeguard of personal freedom — the safeguard of the individual's freedom to decide who they are, what they do and who knows about it [21]. Langheinrich gives a short history of the concept of privacy by design [22] and illustrates as part of that history the origination of five specific categories of privacy that together appear to encapsulate all previous definitions: 1. privacy of personal behaviour (media privacy); 2. privacy of territory (territorial privacy); 3. privacy of the person (bodily privacy); 4. privacy of personal communications (interception privacy); and 5. privacy of personal data (data or information privacy).
Profiling	The extrapolation of information about something, based on known qualities. It leads to the identification of patterns in data of the past which can develop into probabilistic knowledge about individuals, groups and situations in the present and in the future [7].
Risk	A risk is the combination of the chance on and the impact of an undesirable situation. A risk is caused by the combination of an asset, a threat and vulnerability.
Safety and security	Safety is the absence of risk. Security is the absence of risk intentionally caused by others.
Scene	The scene is what can be directly observed by the surveillance system.
Sensor	A device which converts one energy into another, usually an electric signal, e.g. microphone, CCTV camera, pressure sensor or the human eye. There are several closely related concepts. • An active sensor sends a signal which is reflected by the subject and/or which triggers a response from the subject, e.g. radar, sonar and lidar. • An intelligent sensor applies some form of knowledge either to improve the output signal or to interpret the signal to a higher level of abstraction, e.g. a facial recognition system, video analytics or a human. • A probing sensor is a sensor with a probing mechanism with the function of bringing a stimulus to the observed subject. The response to the stimulus is measured by the sensor. Human surveillance professionals do this in security questioning, for example. • A virtual sensor is a sensor in the digital domain, e.g. a sensor that detects a hacking attempt. This is formally not a sensor but typically a software module.
Situational awareness	Situational awareness is the perception of the environment with respect to time and/or space and the comprehension of its meaning. It also includes the projection of the environment into the future or

Term	Definition
	the past.
Stimulus	A stimulus is a detectable change (as perceived by the subject) in the environment (including the subject's own body). A stimulus can already be present in the environment (with the subject passing by) or it can be introduced directly or indirectly by the supervisor. Varying stimuli are used in security questioning and predictive behaviour profiling to trigger a tell-tale reaction.
Subject	In this report, the person under surveillance.
Surveillance	The focused, systematic and routine attention to personal detail for purpose of influence, management, protection or direction [23]. In the context of ERNCIP, surveillance is only covered when used for safety and security purposes.
System	A construct or collection of different elements that together produce results not obtainable by the elements alone [24].
Threat	(That which leads to) the potential occurrence of an undesirable situation. Security measures protect against threats.
Threat assessment	The threat assessment is the process which uses the situational awareness to estimate the concrete threat.
Video analytics	(or video content analysis - VCA) is processing a video to determine spatial and temporal aspects of and relations between objects in a scene. Threat assessment, i.e. generating alerts, is in this report considered to be a separate process.
Vulnerability	A weakness or hole in the security.

Abbreviation	Full text
CI	critical infrastructure
CIP	critical infrastructure protection
ERNCIP	European reference network for critical infrastructure protection
EU	European Union
i-LIDS	image library for intelligent detection systems
MA	morphological analysis
MAS	morphological analysis on the surveillance domain
MAVA	morphological analysis on the subdomain of video analytics
TACTICS	EU KP7 Project: Tactical approach to counter-terrorists in cities
TG (VAS)	thematic group (video analytics and surveillance)
TRL	technology readiness level

List of figures

List of tables

Appendix A Profiling

Profiling is an extrapolation of a characteristic of a person, a group or a situation on the basis of other characteristics of the respective subject. Profiling neither measures nor observes; it is a statistically founded assumption and can therefore never be used as evidence or give weight to other evidence [25]. For example, if profiling (of groups) is used in riot control to decide upon the use of violence, both the chance and the impact of errors in judgement are increased.

Profiling can be characterised in various ways on the basis of time in relation to the incident, input or output variables, object of profiling and application domain. Table 6 provides a list of those characteristics illustrated with examples.

Well known examples of profiling are:

- Criminal geographic profiling, i.e. to find out where the likely perpetrator lives, works or travels.
- Predictive policing is a form of predictive situational profiling, i.e. the profiling of situations or of neighbourhoods.

Profiling is more elaborately described in Section 6.2 of [7].

Table 6 Profiling characterisations

Profiling characterisation	Example	Description
Pre- or post-incident	Predictive profiling	Ascertain the possibility of someone becoming involved in a future incident (as offender).
	Offender profiling; criminal profiling	Ascertain the possibility of someone being involved in an actual incident (as offender) or draw up a profile of the offender.
Input of profiling	Behavioural profiling	Ascertain an aspect of a person (such as their intention) on the basis of their behaviour.
	Racial profiling	Ascertain an aspect of a person on the basis of their ethnicity.
Output of profiling	Geographic profiling	Ascertain a person's residence or place of work on the basis of other aspects.
Domain	Cybercrime profiling	Profiling people or situations in order to prevent or solve cybercrime.
Object of profiling	Person	Profiling people.
	Group	Profiling groups of people (in crowds).
	Situation	Determining whether a situation is suspicious.
	Object	Determining whether e.g. a place is suspicious.

Appendix B Relevant factors for surveillance systems

This appendix contains a description of all dimensions and values of the MA for surveillance systems.

B.1 Context, environment, asset and risk

B.1.1 Context

The context contains all factors that influence the environment but are not actually part of it.

B.1.1.1 Weather

Weather influences visibility within the scene and affects the performance of sensors. It may also influence the behaviour of people and vehicles.

Values are: rain, clear, snowfall, fog (including smog and smoke) and overcast.

B.1.1.2 Weather dynamics

The dynamics of weather can generate movement patterns which may interfere with the function of a surveillance pattern. Highly dynamic weather changes may even generate safety threats.

Values are: stable, slow changing and fast changing.

B.1.1.3 Privacy awareness

Privacy awareness is a mix of the attitude of the people under surveillance and the state of local privacy laws. This is relevant to be able to determine the potential support for surveillance systems.

Values are: none, low (fragmented privacy laws), medium (limited privacy law) and high (integrated and complete privacy law).

B.1.1.4 Security awareness

Security awareness is relevant to determine potential support for surveillance systems. It also typically correlates with the value of assets and perceived risk.

Values are: none, low, medium and high.

B.1.1.5 Intent

High variability of intent of the people in a scene generates highly variable behaviour. The strength of their focus determines whether they can be easily influenced or detracted, which influences the functioning of surveillance capabilities.

Values are: homogeneous and low intensity (e.g. waiting for a bus at a bus stop), homogeneous and high intensity (e.g. waiting in a queue for a security check at an airport), heterogeneous and low intensity (e.g. shoppers in a shopping area) and heterogeneous and high intensity (e.g. train platform once a train has arrived).

B.1.1.6 Relation

The relation between the owner of the object and the owner of the surveillance system is relevant because if the surveillance system is a guest at the location, then the degree to which the environment can be changed to facilitate surveillance is influenced.

Values are: the same (high level of trust and easy-to-change environment) and separate (low level of trust and difficult-to-change environment).

B.1.2 Environment

The environment is all factors that directly interact with the threat, the asset or the surveillance system.

B.1.2.1 Type of environment

This dimension groups several aspects such as indoor/outdoor, the natural environment and the degree to which humans have shaped the environment. This is relevant because structure and stability in the scene reduces the challenge for the surveillance system.

Values are: built high-rise, built low-rise, road, rural, mountains, water (sea, river), underwater, coast, forest, closed indoor (no external light) and open indoor (outdoor lighting).

B.1.2.2 Type of object

The type of object that is under surveillance — or where the asset resides — is relevant for situational awareness.

Values are: compound, house, factory, long infrastructure (pipeline, railroad), flat, bungalow, palace, apartment, bunker, hotel, hospital, public transport hub, street and vehicle.

B.1.2.3 Existing infrastructure

Surveillance systems themselves depend on other infrastructure. The level of infrastructure also correlates with certain threats. Some kinds of infrastructure can be used by the threat or even be a target of the threat.

Values are: none (people illiterate), low (roads, water, electricity and low education), medium (fibre, rail and high education) and high (widespread high-speed internet, divers sensor networks and readily available specialised personnel).

B.1.2.4 Compartments present

The presence of compartments matters for the threat, the situational awareness and the threat assessment. Section 2.6 contains more details.

Values are: vital area (Ring 1), access door/gate, perimeter, secured area (Ring 2) and observation area (Ring 3).

B.1.2.5 Closed compartments

Some compartments may be open for public with only a mild form of access control (e.g. blacklisting). This is relevant for the threat and the situational awareness. The observation area is considered open by definition, so it is excluded here.

Values are: none, vital area (Ring 1), access door/gate, perimeter and secured area (Ring 2).

B.1.2.6 People density

A high density of people attracts different kinds of threats, whereas a low density of people allows for different kinds of surveillance capabilities. A high density of people may require different information demands (e.g. about crowds).

Values are: none, low (< 0.01 person/m^2), medium and high (> 0.5 person/m^2).

B.1.3 Risk

A risk is part of the environment, has a cause and consists of chance and impact.

B.1.3.1 Risk cause

A risk is caused by a combination of asset (target), threat and vulnerability.

B.1.3.1.1 Asset to protect

The asset to protect is relevant for the situational awareness and influences the kind of threat that is directed towards it. However, an actual threat may be targeted towards something else, which is reflected in 'target'.

Values are: public order, the life and wellbeing of a person (e.g. a VIP), the wellbeing of a crowd, the continuation of a process and the integrity of an object (e.g. a high-value transport or a building or technical infrastructure).

B.1.3.1.2 Threat

A threat consists of many elements which may be directly or indirectly observable in the scene.

B.1.3.1.2.1 Target

The target is relevant for the situational awareness and threat assessment. Note that the security system itself may also be a target of the threat. The target type may be other than what the surveillance system was put in place for. A threat can also not be directed at a specific target but be of general presence, e.g. an accident or a weather incident.

Values are: none, VIP, individual, building, monument, crowd, small object (e.g. wallet), ICT infrastructure, industrial infrastructure, vehicle, security system and data.

B.1.3.1.2.2 Threat direction

The threat direction is relevant for the threat assessment and typically determines who has the capabilities or is even responsible for mitigating the threat.

Values are: accident, nature (weather, flooding or earthquake), media, illness (mental), activism, crime, extremism, terrorism and nation state.

B.1.3.1.2.3 Motivation

The threat motivation is relevant for the threat assessment. There may be some empirical correlation with the target type. Nature and accidents do not have a motivation.

Values are: none, political, religious, economic and personal.

B.1.3.1.2.4 Frequency

The frequency (or duration or prior probability) is relevant for the surveillance system. A threat with one short occurrence may be countered with a dynamic mobile surveillance system, while more frequent or persistent threats may require more fixed capabilities.

Values are: single (short), multiple (limited time) and persistent (continuous or no end in sight).

B.1.3.1.2.5 Number of attackers

The number of people that causes the threat is relevant for the situational awareness and the threat assessment. Some threats are not caused by people, but by accidents or nature.

Values are: none, individual, group and crowd.

B.1.3.1.2.6 Capabilities

The capabilities that the attackers have — if any — are relevant for the surveillance system because they can be detected and are valuable for the threat assessment.

Values are: none, finance, weapons (e.g. a gun or explosives), skills (e.g. hacking), capacity (e.g. to quickly grow in number), knowledge (e.g. passwords or how to gain entry) and materials (e.g. tools or camouflage).

B.1.3.1.2.7 Physical angle of attack

The physical angle of the attack is relevant because it determines what kinds of modalities and therefore sensors are needed in the surveillance system.

Values are: ground, air, water surface (e.g. a boat), water subsurface (e.g. a submarine), underground (e.g. through a tunnel) and cyber (through ICT networks).

B.1.3.1.2.8 Modus operandi

The modus operandi is relevant because it determines where the surveillance system should be focused and what kind of threat assessment it must be able to make.

Values are: negligence, sniper, robbery, bombing, suicide attack, poisoning, kidnapping, hostage, burglary, demonstration (riot), molest, vandalism, hacking, infiltration, espionage and theft.

B.1.3.1.2.9 Equipment

The equipment is relevant for the threat assessment and for the required detection capabilities.

Values are: none, explosive (e.g. grenade or IED), tool (e.g. cutter or hammer), poison, banner, UAV, camouflage, electronic device, vehicle, sniper rifle, automatic guns, hand gun, dirty bomb, nuclear bomb and chemical weapon.

B.1.3.1.2.10 Incident phase

The incident phase is relevant for the threat assessment and for the behaviour that the surveillance system needs to be able to observe. The values are a simplification and abstraction of the criminal phases as described in [7].

Values are: before, during and after.

B.1.3.1.3 Vulnerability

If there are no vulnerabilities in the security system (in the broadest sense) then there is no risk, even if there is a threat. Knowing the vulnerabilities is essential for an effective surveillance system. Vulnerabilities result in uncovered attack angles.

Values are: false alarms, intelligence, access control, security awareness, sensor coverage, education, risk assessment and maintenance.

B.1.3.2 Resulting risk

The resulting risk consists of an impact and a chance of that impact. The responsibility of mitigating the risk lies at a certain level.

B.1.3.2.1 Chance

The chance is relevant in terms of prior probability (what is the chance of an attempt). This is relevant because it determines where the focus of the surveillance system should be.

Values are: none, low, medium, high and fact.

B.1.3.2.2 Impact

The impact is relevant because it determines where the focus of the surveillance system should be.

Values are: none, low (people are somewhat affected), medium (people are significantly affected) and high (people are seriously affected).

B.1.3.2.3 Responsibility

The responsibility is relevant because it determines the scope of and support for a surveillance system.

Values are: none, individual, private industry (local company or holdings), regional (community), public (city, province or country), national, supranational (bilateral, EU) and international (UN, NATO).

B.2 Surveillance system: sensors, situational awareness and threat assessment

B.2.1 Sensor

A sensor creates an electrical signal based on what it detects. The sensor is a vital component of a surveillance system. There are many different kinds of sensors and ways to integrate them into a surveillance system.

B.2.1.1 Modality

The modality of a sensor is the physical quantity that it can observe. This is relevant for a surveillance system because it determines what the surveillance system can observe.

Values are: visible light, sound, heat, radio waves, vibration, movement, X-ray, smell and weight.

B.2.1.2 Sensor type

The sensor type is relevant because it determines the modality that can be observed and therefore the objects and attributes in the scene.

Values are: radar, camera, IR-detector, microphone, GSM sniffer, vibration sensor, human, laser, GPS, X-ray sensor, sniffer, sonar, scale and animal.

B.2.1.3 Active

The output of some sensors may improve if the scene is 'lit up' in some modality. Some sensors even require such an active signal (e.g. radar). This is also relevant because it may be observed by people in the scene and can therefore influence them.

Values are: passive, visible light (lamp), heat (infrared), sound (sonar) and radio waves (radar).

B.2.1.4 Invasiveness

The invasiveness of a surveillance system can consist of several aspects [ref]. In this report we focus on the combination of two elements: the degree of cooperation that a subject must give to the surveillance system and the level of detail of information that a surveillance system observes about people in the scene. Both are relevant because they may influence people in the scene.

Values are: not significant, slight (show ability to observe), medium (show ability to intervene) and strong (force a reaction: speaking or standing in the way).

B.2.1.5 Array form

Some virtual sensors are actually composed of multiple sensors or recordings. This matters for the quality of the situational awareness.

Values are: single, stereo, wide baseline, movement and array.

B.2.1.6 Platform

The platform determines the mobility of sensors. This is relevant for the invasiveness and for the quality of the situational awareness.

Values are: fixed, rotating (pan-tilt zoom), limited moving (car) and free moving (animal, human or UAV).

B.2.1.7 Amount of sensors

The amount of sensors available in the surveillance system is relevant for the quality of the situational awareness and for the invasiveness. If there are two or more sensors, then they can be arranged in an array form.

Values are: 1, 2, 10, 100, 1000 and 1 million.

B.2.1.8 Distance sensor–object

The distance between the object(s) in the scene and the sensor determines the quality of the situational awareness and affects the invasiveness as well.

Values are: 0m, 1 m, 10 m, 100 m, 1000 m, 10 000 m. These should be interpreted as ranges of distances (i.e. 1 m implies > 0 m to ~5 m).

B.2.2 Situational awareness

Situational awareness is the perception of the environment with respect to time and/or space and the comprehension of its meaning. It also includes the projection of the environment into the future or the past.

B.2.2.1 Type of object

The type of object that needs to be observed is relevant for the sensor and for the threat assessment. It also relates to the environment and the threat.

Values are: individual, group, letter, parcel, container, bicycle, motorcycle, car, van, bus, tractor, lorry, ship, aircraft, train/tram, heavy machinery, tanker, public service vehicle, weapon and UAV.

B.2.2.2 Type of material to be observed

The type of material is relevant for the sensor and relates to the environment and the threat. For example, wood, cloth (banners) and ceramics (ceramic knife) have to be detected in different manners other than metal, plastics, biological materials or fluids.

Values are: biological, metal, plastic, ceramics, wood, fluid, gas and cloth.

B.2.2.3 Behaviour to be observed

The behaviour in the scene is relevant because it relates to the threat and environment and to the threat assessment and sensor. Detecting loitering is different from detecting trespassing or tailgating.

Values are: loitering, tailgating, trespassing and pickpocketing.

B.2.2.4 Amount of objects to be observed

The amount of objects to be observed is the number of entities of the same kind in the entire scene, including those which do not constitute a threat. This is relevant for the sensors and relates to the environment and the threat.

Values are: 1, 2, 10, 100 and 10 000. These should be interpreted as ranges.

B.2.2.5 Function

The function describes what the surveillance system uses the immediate output of sensors (the signal) for, i.e. how it creates situational awareness. It is relevant for both the sensor and the threat assessment. It is based on the five basic functions of a CCTV camera as described by the United Kingdom Home Office [26], but can also be

generalised for other sensors. More specific functionalities are considered variants of one of these, e.g. tracking is a variant of detection.

Values are: observe, detect, classify and identify.

B.2.2.6 Surveillance pattern

The surveillance pattern is the generic way in which the function is accomplished [8]. It is relevant for the accuracy of the threat assessment and the TRL level.

Values are: threshold alarm, bag of words, concentric circles of protection, profiling and scenario view.

B.2.2.7 Aspect

'Aspect' contains relevant, potential characteristics of an entity. It is not necessary that every entity always has the same aspects: they might even change over time. For instance, an individual always has a face, an identity and is biological, but an individual is not always carrying a weapon, an explosive or is busy communicating. So there may be a link between those two dimensions (entity and aspect) and the actual presence depends on the situation and circumstances. It is relevant for the sensor and for the threat assessment.

Values are: presence, number plate, face, weapon, behaviour, communication, identity, tools, electronic device, explosive and material.

B.2.2.8 Accuracy

The required accuracy is very relevant for the threat assessment and the sensor. It also relates to the environment and the threat. We have excluded 100 % accuracy because it is wise to always take potential errors into account.

Values are: 20 %, 50 %, 90 %, 99 % and 99.99%.

B.2.3 Threat assessment

The threat assessment is the process which uses the situational awareness to estimate the concrete threat. The difference between the factors grouped under 'threat' and those grouped under 'threat assessment' is that the former constitute the actual threat, while the latter constitute what the surveillance system needs to know about the threat. The first is something that happens to the surveillance system and the second is something that the surveillance system is explicitly designed to deliver.

B.2.3.1 Security process

The security process relates to the threat because some threats are easy to prevent with a surveillance system but hard to investigate, and vice versa. It also creates the information demand that must be addressed by the situational awareness function. These are introduced in Section 2.2.

Values are: preparation, intelligence, prevention, in-the-act and investigation.

B.2.3.2 Threat assessment

The threat assessment is relevant because it determines what is needed from the situational awareness. It is also related to the actual threat.

Values are: threat direction, threat motivation, frequency, number of attackers, capabilities, threatening objects or persons, modus operandi, equipment, target, physical angle of attack and incident phase.

B.2.3.3 Reliability threat assessment

The reliability of the threat assessment is relevant because it is related to the situational awareness, the actual threat and the asset to protect.

Values are: 10 %, 50 %, 90 %, 99 %.

B.2.4 System

The surveillance system also has some generic factors because it is a system. Like any complex system, there are development phases and technology readiness levels to consider. Cost could also be a factor here, but that is left for others to reflect upon.

B.2.4.1 Development phase

This is relevant because different functionalities are required in different phases of the development. For example, automatic calibration could be useful in the configuration and maintenance phase, but no longer in the use phase.

Values are: installation, configuration, use, maintenance and decommission.

B.2.4.2 Technology readiness level

The required TRL is relevant for the performance and many other factors. Much technology has been developed that has not yet reached TRL 9 but that may still be very useful on a short notice.

Values are: TRL 1-9.

B.2.4.3 Subcomponents

The composition of a surveillance system varies depending on the function. A minimal surveillance 'system' consists of a person, but typically many other components are also required. The composition includes various subcomponents which have consequences for the quality of the threat assessment and other factors.

Values are: sensor, storage, network, processing unit, viewing station, mobile interface, command and control unit, human operator and human on the floor.

Appendix C Relevant factors for video analytics

This appendix contains a description of all dimensions and values of the MA for video analytics.

C.1 Scene

The scene is everything that can be observed directly through a camera. Video is not well suited for observation in high precipitation or foggy (smog or smoke) conditions.

C.1.1 Cover

The coverage of the scene is relevant because weather and the day/night cycle may influence the quality of the video analytics.

Values are: indoor and outdoor.

C.1.2 Light

Depending on the relative placement to the camera, additional lighting on a scene can help to increase visibility. This is a specialisation of the dimension 'active' of a surveillance system.

If the camera is sensitive in the infrared spectrum then it is an option to light only in infrared. This may prevent detection of the surveillance system.

Values are: none, lamp and infrared.

C.1.3 Amount of objects

The amount of objects in a scene is relevant because they may occlude each other and thereby prevent a clear view.

Values are: 1, 5, 10, 100 and 1000.

C.2 Camera

The camera converts light into an electrical (digital) signal. Many aspects of the camera are highly relevant for the quality of video analytics because if certain information is not reflected in the signal, then no amount of video processing can retrieve it again.

In this section we consider an array of cameras (e.g. stereovision) as one signal source.

C.2.1 Distance to object

The distance between the object(s) in the scene and the camera determines the quality of the situational awareness.

Values are: 0 m, 1 m, 10 m, 100 m, 1000 m. These should be interpreted as ranges of distances (i.e. 1 m implies > 0 m to ~5 m).

C.2.2 Orientation

The orientation of the camera to the scene is relevant because it can influence the amount of occlusion. It is also relevant because certain aspects of objects can only be observed from specific angels (e.g. a face cannot be seen from above).

Values are: top down, skim and side.

C.2.3 Modality

The modality is relevant because some give more detail, while others may be invariant to the day/night cycle. The type of lighting should be adapted to the required modality.

Values are: visible light and heat.

C.2.4 Array form

The array form is relevant because it may help to avoid occlusion and may give additional information about the distance between the object and the sensor or the angle.

Values are: single, stereo and wide baseline.

C.2.5 Platform

The platform to which the camera is attached is relevant because it determines the mobility of the camera, which in turn influences the array form, the distance to the object and the orientation of the camera.

Values are: fixed, rotating (pan-tilt zoom camera), limited moving (e.g. mounted on a car) and free moving (e.g. a handheld mobile phone or a UAV).

C.3 Video processing chain

The video processing chain uses video to determine spatial and temporal aspects of objects in the scene. It also includes steps to improve the video with regard to the purpose of the system.

C.3.1 Video signal

The video signal can be improved by reducing the size and therefore the required bandwidth and/or by somehow improving the quality of the video signal.

C.3.1.1 Image improvement

Image improvement is relevant because it may enhance certain useful aspects of the video, but this will typically come at the cost of others. In this report, it is considered different from video analytics because the function is not the same, i.e. the output of this step is still an image. Separating them allows for more specificity in these descriptions and their relations.

Values are: masking, stitching, super resolution and stabilisation.

C.3.1.2 Compression

The amount of compression is relevant because it may introduce artefacts into the signal which negatively influence the quality of the situational awareness. Certain mobile platforms (e.g. UAV or mobile phone) may only support a low-bandwidth data link which implies that compression is required.

Values are: none, low and high.

C.3.2 Video analytics

Video analytics uses video to determine spatial and temporal aspects of and relations between objects in the scene.

C.3.2.1 Function

The function of video analytics is relevant because it directly determines the situational awareness and is relevant for the sensor and the required video signal. The function should be matched with the scene, and sometimes adaptations will be made in the scene to facilitate this function.

While there is no specific order in these functions, they can roughly be arranged by the amount of detail of information they generate on an object.

Values are: observation, detection, tracking, calibration (i.e. obtaining the orientation of an object), shape recognition, classification, recognition (equals identification out of a limited set) and identification (equals recognition out of all objects).

C.3.2.2 Technology readiness level

The required TRL is relevant for the performance and many other factors. Much technology has been developed which has not yet reached TRL 9 but may still be very useful on short notice.

Values are: TRL 1-9.

C.4 Situational awareness

Situational awareness is the perception of the environment with respect to time and/or space and the comprehension of its meaning. It also includes the projection of the environment into the future or the past, but excludes higher level information such as the modus operandi. This is considered in the MAS.

C.4.1 Object type

The object type is relevant because it influences the kind of video processing chain that is required.

Values are: person, vehicle, camera, zone, baggage, doorway, crowd and fence.

C.4.2 Aspect

The aspect of an object is relevant because it determines the kind of video processing chain that is required.

Values are: presence, location, flow, identity, behaviour (including interaction), orientation (e.g. the calibration of a camera), class and quantity.

C.4.3 Relation

The relation between two objects that is to be determined is relevant because it determines the kind of video processing chain that is required.

Values are: temporal, spatial, causation, legal and interaction.

C.4.4 Accuracy

The accuracy that is required is relevant because it is directly related to the TRL level and the required functionality of the video processing chain.

Values are: 50 %, 90 %, 99 % and 99.99 %.

Appendix D List of relevant factors for surveillance systems

This appendix contains a flat list of all dimensions and values of the MAS. This is convenient for quickly generating new configurations.

Weather: rain, clear, snowfall, fog (including smog and smoke) and overcast;

Weather dynamics: stable, slow changing and fast changing;

Privacy awareness: none, low, medium and high;

Security awareness: none, low, medium and high;

Intent: homogeneous and low intensity, homogeneous and high intensity, heterogeneous and low intensity and heterogeneous and high intensity;

Relation: the same and separate;

Type of environment: built high-rise, built low-rise, road, rural, mountains, water, underwater, coast, forest, closed indoor and open indoor;

Type of object: compound, house, factory, long infrastructure, flat, bungalow, palace, apartment, bunker, hotel, hospital, public transport hub, street and vehicle;

Existing infrastructure: none, low, medium and high;

Compartments present: vital area (Ring 1), access door/gate, perimeter, secured area (Ring 2) and observation area (Ring 3);

Closed compartments: none, vital area (Ring 1), access door/gate, perimeter and secured area (Ring 2);

People density: none, low, medium and high;

Asset to protect: public order, the life and wellbeing of a person, the wellbeing of a crowd, the continuation of a process and the integrity of an object;

Target: none, VIP, individual, building, monument, crowd, small object, ICT infrastructure, industrial infrastructure, vehicle, security system and data;

Threat direction: accident, nature, media, illness, activist, criminal, extremist, terrorist and nation state;

Motivation: none, political, religious, economic and personal;

Frequency: single, multiple and persistent;

Number of attackers: none, individual, group and crowd;

Capabilities: none, finance, weapons, skill, capacity, knowledge and materials;

Physical angle of attack: ground, air, water surface, water subsurface, underground and cyber;

Modus operandi: negligence, sniper, robbery, bombing, suicide attack, poisoning, kidnapping, hostage, burglary, demonstration, molest, vandalism, hacking, infiltration, espionage and theft;

Equipment: none, explosive, tool, poison, banner, UAV, camouflage, electronic device, vehicle, sniper rifle, automatic guns, hand gun, dirty bomb, nuclear bomb and chemical weapon.

Incident phase: before, during and after;

Vulnerability: false alarms, intelligence, access control, security awareness, sensor coverage, education, risk assessment and maintenance;

Chance: none, low, medium, high and fact;

Impact: none, low, medium and high;

Responsibility: none, individual, private industry, regional, public, national, supranational and international;

Modality: visible light, sound, heat, radio waves, vibration, movement, X-ray, smell and weight;

Sensor type: radar, camera, IR-detector, microphone, GSM sniffer, vibration sensor, human, laser, GPS, X-ray sensor, sniffer, sonar, scale and animal;

Active: passive, visible light, heat, sound and radio waves;

Invasiveness: not significant, slight, medium and strong;

Array form: single, stereo, wide baseline, movement and array;

Platform: fixed, rotating, limited moving and free moving;

Amount of sensors: 1, 2, 10, 100, 1 000 and 1 million;

Distance sensor–object: 0 m, 1 m, 10 m, 100 m, 1 000 m, 10 000 m;

Type of object: individual, group, letter, parcel, container, bicycle, motorcycle, car, van, bus, tractor, lorry, ship, aircraft, train/tram, heavy machinery, tanker, public service vehicle, weapon and UAV;

Type of material to be observed: biological, metal, plastic, ceramics, wood, fluid, gas and cloth;

Behaviour to be observed: loitering, tailgating, trespassing and pickpocketing;

Amount of objects to be observed: 1, 2, 10, 100 and 10 000;

Function: observe, detect, classify and identify;

Surveillance pattern: threshold alarm, bag of words, concentric circles of protection, profiling and scenario view;

Aspect: presence, number plate, face, weapon, behaviour, communication, identity, tools, electronic device, explosive and material;

Accuracy: 20 %, 50 %, 90 %, 99 % and 99.99 %;

Security process: preparation, intelligence, prevention, in the act and investigation;

Threat assessment: threat direction, threat motivation, frequency, number of attackers, capabilities, threatening objects or persons, modus operandi, equipment, target, physical angle of attack and incident phase;

Reliability threat assessment: 10 %, 50 %, 90 % and 99 %;

Development phase: installation, configuration, use, maintenance and decommission;

TRL: 1, 2, 3, 4, 5, 6, 7, 8 and 9;

Subcomponents: sensor, storage, network, processing unit, viewing station, mobile interface, command and control unit, human operator and human on the floor;

Appendix E List of relevant factors for video analytics

This appendix contains a flat list of all dimensions and values of the MAVA. This is convenient for quickly generating new configurations.

Cover: indoor and outdoor;

Light: none, lamp and infrared;

Amount of objects: 1, 5, 10, 100 and 1 000;

Distance to object: 0 m, 1 m, 10 m, 100 m, 1 000 m;

Orientation: top down, skim and side;

Modality: visible light and heat;

Array form: single, stereo and wide baseline;

Platform: fixed, rotating, limited moving and free moving;

Image improvement: masking, stitching, super resolution and stabilisation;

Compression: none, low and high;

Function: observation, detection, tracking, calibration, shape recognition, classification, recognition and identification;

TRL: 1-9;

Object type: person, vehicle, camera, zone, baggage, doorway, crowd and fence;

Aspect: presence, location, flow, identity, behaviour, orientation, class and quantity;

Relation: temporal, spatial, causation, legal and interaction;

Accuracy: 50 %, 90 %, 99 % and 99.99 %;

Appendix F Auto-calibration

In the context of surveillance and specifically of cameras, auto-calibration is a (partly) automated process to determine metadata about one or more sensors, such as a camera, from the data of the respective sensor itself. This includes its position and orientation, but can also be other properties such as sensitivity, colour balance, etc. It can be used efficiently, because automated, to:

— deploy and maintain a large-scale multisensor network, such as a surveillance system with cameras;

— filter sensors which are suitable for a specific capability (e.g. the tracking of persons);

— assess which capabilities are possible given a specific set of sensors.

Auto-calibration is generally still a topic of research [27]. The closest functionality may be 'tampering detection', which is the automatic detection if some properties of a sensor have been changed.

The MAVA is as follows.

MAVA — Function: calibration; object type: camera; aspect: location and orientation.

A special type of auto-calibration is ego-motion estimation (EME). EME estimates the movement of the sensor platform on its output data. This can be useful with UAVs, sensors on vehicles, body cams and wearables.

F.1 In which scenarios is auto-calibration useful?

Auto-calibration is also very convenient in large-scale CCTV deployments where maintenance and moving threat patterns lead to continuous changes in the CCTV conditions. Consider a more specific scenario where one organisation is responsible for the deployment and maintenance of sensors and another is responsible for the capabilities created with the sensors, i.e. it uses them. Auto-calibration can help to efficiently and objectively assess the current status of the sensors, which can facilitate the cooperation between these parties.

It also helps in the development of new algorithms to stay in control of the conditions under which the new type of analytics is being deployed in the field. This helps prevent teething troubles and manage clients' expectations.

In the case of a crisis, such as a terrorist attack, LEAs suddenly gain (legal) access to many more surveillance cameras they do not work with under normal circumstances. Under normal circumstances they are used to sustain a certain type of capability and therefore may be equipped with different types of analytics — or none at all — than are required during a crisis. Auto-calibration can help to filter cameras that are suitable for a certain capability, and to filter capabilities that are suitable for a certain camera configuration.

F.2 What are the alternatives to auto-calibration?

There are two alternatives to auto-calibration. The first is manual calibration, which may require introducing objects onto the scene, such as large checker-boards with patterns or a Rotakin doll, as well as manually annotating certain points or lines in the footage. If done properly, this method can yield the most accurate results, but visiting all relevant scenes may not be economically feasible and sometimes even dangerous or impossible because the sensor has already left the environment.

The second alternative is:

(1) to make the system so robust that any kind of use will not result in unreliable system properties;

(2) to (re)configure the system with the required specificity for all future use cases;
(3) to store the initial properties and make them available through standard protocols (e.g. ONVIF or PSIA [28]).

The problem with only applying this method is that the first two steps are not very realistic for larger surveillance systems in the context of security. Camera properties may drift and larger systems are prone to changes due to maintenance. In addition, it is rather difficult to foresee all future use cases, e.g. due to moving threat patterns, that arise from crises. Finally, the use cases that are relevant during crises will typically be more demanding than those of regular use cases. More demanding use cases typically imply that more system properties must be known. The calibration information that is sufficient for regular use will therefore typically not be sufficient during crises.

In practice, a combination of these three methods may be required.

F.3 What is the difference between (re)configuration, justification and gauging?

The configuration of a system is the arrangement of elements, including possibly sensors, in a particular form. It can include both the physical (hardware) and logical (software) elements. A system always has an initial configuration. To configure a system is to determine its initial state. To reconfigure it is to adapt the initial configuration. A justification is a special kind of configuration, which is to change the configuration to a predetermined standard. The system has to be changed for there to be configuration.

To calibrate a system is only to determine certain aspects of that system. It is only necessary to observe (read or measure) the system for calibration.

To gauge is to combine calibration and justification together.

F.4 What is the difference between auto-calibration and making a 3D model of an environment?

Calibration only concerns estimating the properties of the sensors, which could also be in groups or in three dimensions. A 3D model of the environment also includes the buildings, vehicles, etc. and, in most cases (depending on the definition of 'environment'), it excludes the calibration data.

The two processes, '3D modelling' (or 3D reconstruction) and calibration, are closely linked. On the one hand, it is impossible to use its output data to create a 3D model without information about the properties of the sensor. On the other hand, the calibration process can be simplified with the background knowledge of a 3D model of the sensor's environment.

F.5 Why is auto-calibration difficult?

With auto-calibration it is attempted to estimate sensor properties from the output data of the respective sensors itself. This requires background knowledge of the environment and/or the sensor. For example, are the moving objects in the scene pedestrians or cars? For optimal results, auto-calibration must start with the right type and amount of background knowledge without making erroneous assumptions.

It also requires sufficiently rich data: if useful data exist only in a small corner of the frame and only by night, then the quality of the resulting calibration will not be optimal.

Next, small deviations in the sensor properties can have great consequences for the quality of the outcome. This is especially true for automated processing steps, as they are typically less forgiving of such deviations than human operators.

Finally, the number of sensor properties relevant for useful applications and the number of interdependencies between them are both large: is the camera zoomed in or is it closer to the scene?

As of the end of 2015, there are no large data sets available for testing and developing auto-calibration.

Europe Direct is a service to help you find answers to your questions about the European Union
Free phone number (*): 00 800 6 7 8 9 10 11
(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the internet.
It can be accessed through the Europa server http://europa.eu

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

LB-NA-27852-EN-N

JRC mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy directorates-general, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

Publications Office

doi:10.2788/945388

ISBN 978-92-79-57771-0