Context Ontology Modelling for Improving Situation Awareness and Crowd Evacuation from Confined Spaces

Gianluca Correndo, Banafshe Arbab-Zavar, Zlatko Zlatev, Zoheir A. Sabeur

Context Ontology Modelling for Improving Situation Awareness and Crowd Evacuation from Confined Spaces

Gianluca Correndo, Banafshe Arbab-Zavar, Zlatko Zlatev and Zoheir Sabeur

{gc, zdz, baz, zas}@it-innovation.soton.ac.uk
University of Southampton IT Innovation Centre, Electronics and Computer Science, Faculty of Physical Sciences and Engineering, Southampton, United Kingdom.

Abstract. Crowd evacuation management at large venues such as airports, stadiums, cruise ships or metro stations requires the deployment and access to a Common Operational Picture (COP) of the venue, with real-time intelligent contextual interpretation of crowd behaviour. Large CCTV and sensor network feeds all provide important but heterogeneous observations about crowd safety at the venue of interest. Hence, these observations must be critically analyzed and interpreted for supporting security managers of crowd safety at venues. Specifically, the large volume of the generated observations needs to be interpreted in context of the venue operational grounds, crowd-gathering event times and the knowledge on crowd expected behaviour. In this paper, a new context ontology modelling approach is introduced. It is based on knowledge about venue background information, expected crowd behaviours and their manifested features of observations. The aim is to improve situation awareness about crowd safety in crisis management and decision-support.

Keywords: crisis management · decision-support · sensor data · semantics · context ontology modelling · crowd behavior

1 Introduction

Advanced situation awareness, safety of crowd and their evacuation from large mass gathering venues during incidents is of paramount importance in crisis management. The crowd must be evacuated rapidly towards safer zones. However, the real-time monitoring of crowd during incidents is an enduring responsibility taken by security managers. Nowadays security personnel is able to observe venue spaces in real-time through using ICT based devices and use automated systems for their decision-support. The ICT devices and systems may include a large network of deployed CCTV cameras, multiple arrays of sensors and a Common Operational Picture (COP) of the venue of interest. Nevertheless, security managers can be overwhelmed by sheer volumes and complexity of observation data while making their decisions on crowd spatial evacuation in real-time. Further, they could potentially encounter risks of failure in their decisions since they may not be able to capture and interpret incident related situations critically on time.

The eVACUATE1 project specializes in crowd behaviour detection using multiple video observations under the visible, thermal and hyper-spectral light bands, which is also enhanced by high level crowd psychology and typology knowledge reasoning, modelling and fusion. This shall add value to existing crowd evacuation strategies with much improved situation awareness and decision-support. Four crowd evacuation pilot experiments are considered in the project. These include the following: 1- Evacuation of a Soccer Stadium - Real Sociedad de Futbol S.A.D (Spain); 2- Mustering and evacuation of passenger cruise ship -STX-FR (France); 3-Airport Terminal evacuation - Athens International Airport(Greece); and 4- Metro Tube evacuation – Metro Bilbao S.A (Spain).

Crowd behaviour modelling is an established discipline in numerical phenomena simulation. Also, with more affordable and easy to deploy ICT devices to monitor crowd, it has become possible to analyze human behaviour with more degrees of spatial granularities and context to advance knowledge reasoning and crowd behaviour modelling. "Normative behaviour models" are based on established cognitive psychology theories, or "Script" (Schema) theories [1]. These postulate that human behaviour can mostly be described as a set of patterns, so called "scripts". These are basically comparable to the way a written script is set with its provision of a program for actions. For example, the stringent requirements on following safety procedures on board aircrafts or cruise ships is a typical attempt to set up appropriate temporary schema to be used during emergency [2]. Therefore, it is highly desirable to investigate on the scripts which can be generated by crowd during a real evacuation event from a given venue. This will provide a good basis for understanding how crowd evacuation evolves in space and time and also learn from it for further improvement on future evacuation plans. It is worth noting that understanding usual (and unusual) behaviour in context of crowd types and gathering venues is essential to proceed with this study.

Automated usual (and unusual) behaviour detection aims at supporting security managers and decision-makers during normal crowd monitoring activities as well as during evacuation. In this paper, an ontological framework based on causality assumption on crowd behaviour is developed. This is in contrast to the traditional rigid approaches based on the creation of ad-hoc rules which describe crowd behaviour. Moreover, the proposed approach introduces context knowledge as a source of important information, to improve the crowd behaviour classification process and modelling.

2   Context Knowledge and Human Behaviour Detection

Research interest on domain knowledge usage for improving the performance of traditional machine learning based human behaviour detection classifiers from vision observation has been growing over the last few years. Concepts of ambient intelligence which are embedded in cognitive vision systems have also been explored. The

---

use of context about crowd typology in accord with mass-gathering venue characteristics, including operational spaces and times are being researched for advancing the performance of crowd behaviour detection algorithms using vision in the eVACU-ATE project.

2.1   Contextual Knowledge in Cognitive Vision

In [3] it is presented an approach, named *key-frame*, for modelling human sequences within cognitive vision systems that allow the description and synthesis of human actions within spaces. The approach focuses on actions performed by individuals, segmenting them into key frames which are characterised by body posture configurations. At higher level of abstraction, human behaviour(s) are described in a lattice of symbolic templates, named situation graph trees. The graph trees describe the relationships between agents and the surrounding environment. Various types of semantic representations have been proposed within the framework of human sequence evaluation. Depending on the addressed domain of safety and security surveillance, they can support the classification of critical situations accordingly.   Within the surveillance domain, an ontology has been proposed [4], based on the key-frame approach. The core of this ontology is introduced with the following listed concepts: 1) **Event** (or situation): Characterized by a status, which encodes metric-temporal knowledge; 2) **Contextualized Event**: Describes higher level interactions among entities; and 3) **Behaviour Interpretation**: Specifies surveillance oriented behaviors of interest, i.e. "abandoned object". One of the main issues in applying such approach for monitoring crowd in EVACUATE's confined spaces is that it will be almost impossible to maintain the identity of the monitored entities and rendering the tracking of agent's activities through frames. Moreover, there is no design pattern that will support the creation of general behaviour(s). This approach seem to be too rigidly linked to the native domain of crowd surveillance and cannot be proposed as a reusable knowledge for modelling.

2.2   Context in Ambient Intelligence

Ambient intelligence concept ontologies are used to provide high level descriptions of situations in order to support context awareness in smart applications for decision-support. The term "context" is understood to be: "*Any information that can be used to characterise the state of an entity*" [5]. The aim of ambient intelligence is also for recognising what users do so that they can be assisted in spaces around which they move. This aim somehow steps on the definition provided on "human behaviour" [6]. Context here plays a key role for understanding not only the activities performed by agents within a spatial environment, but also how these activities can be interpreted within a broad framework of relationships between other agents, entities and unfolding events. These concepts have been modelled using ontologies such as: COBRA-ONT [7] and SOCAM [8]. Nevertheless, such ontologies provide concepts only for user-oriented entities such as *Location*, *Person*, or *Activity*. They are solely focussed

in personalised ambient intelligence but with less interest on the specificies of human behaviour detection.

3 Context Information in Crowd Evacuation Management

Context in ambient intelligence is understood to capture the amount of information that could characterize a situation of an entity at a given space of interest. It is partly given and/or acquired, while it remains user interactive and application centric. In eVACUATE, context information and sensor observations are kept separate. Nevertheless, Context is used in the interpretation of the sensor observation. With this in mind, context is defined as the amount of *a priori* knowledge information about expected crowd behaviour(s) at given venues of interest. This prior knowledge is assumed to give trade-offs on better interpreting and detecting crowd behaviour from video sensing. During the monitoring of crowd in real-time, it is important to model the expected interactions of crowd(s) at various spatial scales. These include macro-, meso- and micro-scales. They could represent overall people, finite number of people as a group and single individuals within the crowd respectively. These types of interactions can be modelled by representing them as causal connections across *a prior* known behaviour modes, contextual information and observed (detectable) features.

3.1 New Crowd Behaviour Contextual Modelling

Ontological models are used to describe crowd *Usual* (expected) behaviour(s) in context of venues. These models describe how behaviour manifests itself and how it compares with its detectable features through sensor observations. When discrepancies between the expected and detected behaviour occur, then the behaviour is classed as *Unusual*. Equally, when discrepancies do not occur, the behaviour is classed as *Usual*. The distinction between *Usual* and *Unusual* behaviour in this study is mainly dependent on the following: *1) Context; 2) The Model which describes the situation; and 3) The set of observations which are collected from a venue*. Similarly, subdivisions of *Context*, *Model*, and *Observations* are found in logical problem definitions such as *diagnosis* and *diagnostic* in model-based reasoning [9]. For instance, Diagnosis is described with a set of assumptions that are joint with a logical model system (i.e. Crowd at a given Venue) and context to explain observations. For example, when monitoring a flow of spectators at different parts of a venue, like a football stadium during a match, usual behaviour is specifically characterised in accord with the place that is monitored (e.g. *Entrance Gate*; *Ticket Counters*, *Bar*, *Club Fan Shop*, *Stands* etc.). Usual behaviour is also characterized in accord with the various stages of the unfolding event (e.g. *Pre-Match*; *Match in progress*; *Post-match*).

- *Pre-Match*: The spectators are expected to enter the stadium through the *Entrance Gates*. From there they proceed to the *Stands*.
- *Match in progress*: The expected behaviour at the stands is for spectators to be partly stationary on their benches (i.e. with oscillatory or occasional movements of

small groups for toilet breaks and trips to the bar). Hence, any substantive movement of individual or groups of spectators towards opposite fans can be considered as unusual.

- *Post-Match*: Groups of individuals are expected to leave the stadium and head towards the car parking spaces.

The above mentioned example of venue and crowd type clearly shows that the various stages of the football match event together with the various monitored areas around the stadium are by far the most important knowledge for defining the expected (Usual) behaviour of spectators(crowd or fans). Specifically, crowd physical motions related parameters such as the predominant crowd velocities (speed and direction) and densities (counted number of people per unit area) become important to measure. In this case, rules that model expected crowd behaviour at a football match can be simply defined as follows:

$$BehavioralMode(usual) \land Venue(stadium) \land Time(pre\text{-}match) \land Place(gate) \vdash$$
$$Expected(flow(X, gate, benches) \land velocity(X,V) \land V \leq 5m/s) \tag{1}$$

The above logical formula (1) involves the flow velocity V (including speed and direction) of detected groups X which are represented as predicates. It also includes the context information of Venue, Type, and Place. The *BehaviouralMode* term in (2) does not belong to the context; its value (i.e. "usual") is the assumption made by the system when processing observations. Formula (1) represents the expected usual crowd motion conditions (i.e. usual behaviour). However, the spectrum of behaviour(s) can be formulated at various spatial scales of detected groups in the crowd. These concern the generic terms X and their respective estimated velocities. The system can use such formulas by adopting available contextual information and assuming usual behavior(s). Forward logical chaining can be then applied to get a set of expected values of observations. The expected values will be then matched against actual observed features that need to be detected computer vision based machine learning algorithms. (Equally, the system can use the concept of formula (1) for backward chaining observations and context to infer values of predicated behavior.

4 Ontology Design Pattern for Context Definition

In order to correctly interpret crowd sensor observations, a set of ontological behavioural templates needed to be developed against measured sensor properties (or features). The latter are extracted from crowd behaviour detection algorithms trained on vision sensor observation data sources (e.g. video feed). OGC standards on "Sensor Observations and Measurements" [10] were adopted to develop the behavioural templates of eVACUATE. OGC ontologies such as SSN ontology [11] were specifically used to semantically enhance the concept of "Observation" in this research study.

4.1 Context for Defining Crowd Behaviour

Some crowd behaviour(s) are context agnostic (e.g. Crowd stampede). These do not need further contextual information to fully define them. However, most crowd behaviour(s) are context dependent. Their interpretation depend on a number of contextual factors which are described (with examples) below:

- *Time*
  The flow of spectators at a stadium change with the expected match schedules (pre-match, during the match, after match). Or, passengers motions at an underground station platform vary depending on the train time table (before, during, after train arrives)
- *Space*
  Passengers in an airport tend to queue at check-in desks; or at ticket kiosks of an underground station. Also, spectators at a stadium are usually orderly seated at stands.
- *Venue*
  Passengers evacuating from a cruise ship are expected to behave differently from spectators evacuating a stadium.
- *Other*
  Crowd(s) behaviour expectations from instructions (i.e. Audio guided evacuation at a given venue).

The behaviour templates associate the above behaviour-like modes in this study. These have key detectable features which can be observed through sensing hods, including video cameras
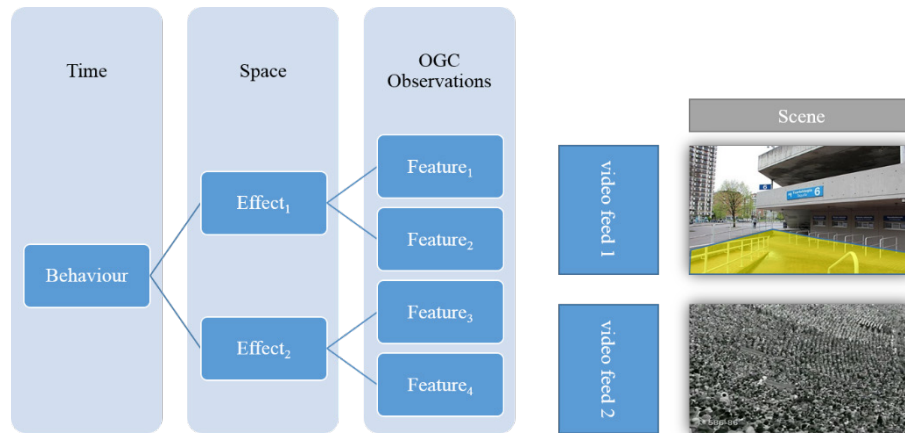

Fig. 1. Behaviour Template

At the heart of the context behaviour ontology describing behaviour templates lies the concept of causality which is used to match the interpretation of the collected sensor observations. The crowd contextual behaviour ontology distinguishes three main components (see Fig. 1): 1-*Behaviours*; 2-*Effects*; and 3-*Features*.

Sensor observations are modeled as the detected manifestations of the effects of unfolding behaviour(s). Crowd *Behaviour*(s) are described in templates of expected human interactions in a given venue. They are defined with a set of expected *Effect*(s) in space and time. *Effect*(s) are spatially tangible (or intangible) result, in terms their manifested detectable features. They are also conceptualized as spatially related to scheduled activities of crowd(s) at the venue. The manifested *Feature*(s) of such *Effect*(s) are sensory observed at the venue with specific measurable values.

OGC O&M standards on "Observation & Measurements" [10] define observation as:

"An Observation is an action whose result is an estimate of the value of some property of the feature-of-interest, at a specific point in time, obtained using a specified procedure".

Although our approach is inspired from OGC O&M standards, the crowd behaviour modelling template was developed distinctly in terms of its description of observations and features. In it, the Effect(s) are postulated as a collection of detectable Feature(s). Also, Feature(s) are constrained by a set of specific observations with detectable properties. For example, sensor readings of property, which concern crowd activities in a venue, may include crowd physical motion (Velocity: Speed and direction); and density (Number of individuals per unit spatial area) at various spatial resolutions and levels. The main distinction from OGC O&M standards in the construction of the behaviour template here, mainly concerns the restriction of the possible (and range of) values of observations of given Feature(s). This is greatly supported by the assumption that crowd(s) are expected to behave in a particular way within a given and venue and times. This new distinction enables good reasoning on observed and detectable Usual (or Unusual) crowd behaviour.

The inclusion of effects with no direct tangible features are required in order to model psychological states that involve single (or group of) individuals in the crowd(s). These are not directly detectable via sensing but undoubtedly influence the value and range of detectable features. In eVACUATE, the modelling of crowd behaviour(s), the spatial environment in which they unfold and their evolution in time, allows the abstraction of behaviour templates and instantiate them generically for different of venues and crowd(s). For example, in the expected behaviour at a stadium, which was introduced in Section 3, can be generalized using an abstract behaviour template (See Table 1) so that it can be used for different venues. (For a Stadium, the gathering places are the stands, the transition places are the gates and the internal corridors, while the exit can be the car park outside the stadium.)

Table 1. Event Behaviour Template

Time				
		pre	**During**	**post**
Space	**Gathering Place**	People settling in	People stably at place	People leaving from the gathering place
	Transition Place	People moving towards gathering place	Mostly empty	People moving towards the exit
	Exit	People moving towards inside the venue	Mostly empty	People moving towards the outside of the venue

4.2 Modelled Crowd(s) Features

The ontological modelling of detectable features is a process which is primarily driven by the type of information that can be collected from the deployed sensor networks at the venue. At this stage of the project, the primary source of information are CCTV cameras which are already installed at the venue. The included features in the modelling so far, include the ones which are available by the feature extraction algorithms in eVACUATE. These particularly focus on venues with the description of crowd(s) at various scales and levels. These are highlighted below:

1. *Micro-scale*: An individual is identified with measurable time-dependent position and velocity
2. *Meso-scale*: Although group of individuals are identified with measurable time-dependent positions and velocities, these are represented as a statistical distribution in addition with other features. These are included as: a) *Density* which is measured as number of people per unit area. This is used to model levels of service as described by Fruin [12] and the expected presence of people within target areas in a venue; and b) *Direction*: A symbolic information which described that a group is intentionally moving towards an explicitly described target (e.g. an exit) with a certain degree of certainty. This is used additionally to model the expected flow of people
3. *Macro-scale*: Average quantities such as crowd density, dominant motions (Translations), structure and internal energies are determined in this scale. They specifically include the following features: a) *Internal Energy*: Normalized between 0 and 1, it estimates the amount of oscillations of elements within a crowd; b) *Structure*: Normalized between 0 and 1, it estimates the strength of the connections within a crowd; and c) *Predominant Translation*: Used to measure the degree of mobility of the crowd.

4.3   Ontology Modules and Sensor Data Ingestion Architecture

As mentioned in Section 4.1, behaviour templates can be generalized and reused for different domains (venues) and sub-classing. The behaviour template in Section 4.1 concerning the "Stadium domain" is further detailed. As depicted in **Fig. 2**, the Event Behaviour Template captures the pre-event period (i.e. before the event takes place) with the expected effect as a Gathering Effect. This effect manifests itself at the Gates and Corridors, while the crowd moves towards the stands. The Crowd direction of motion is depicted with a target → Stadium Stands. As a result, the expected usual crowd(s) Gathering Behaviour at a Stadium (Football match event) states the crowd(s) move without running, while they maintain a high structure. This is expected from groups moving towards shared destinations and not scattering around the venue.
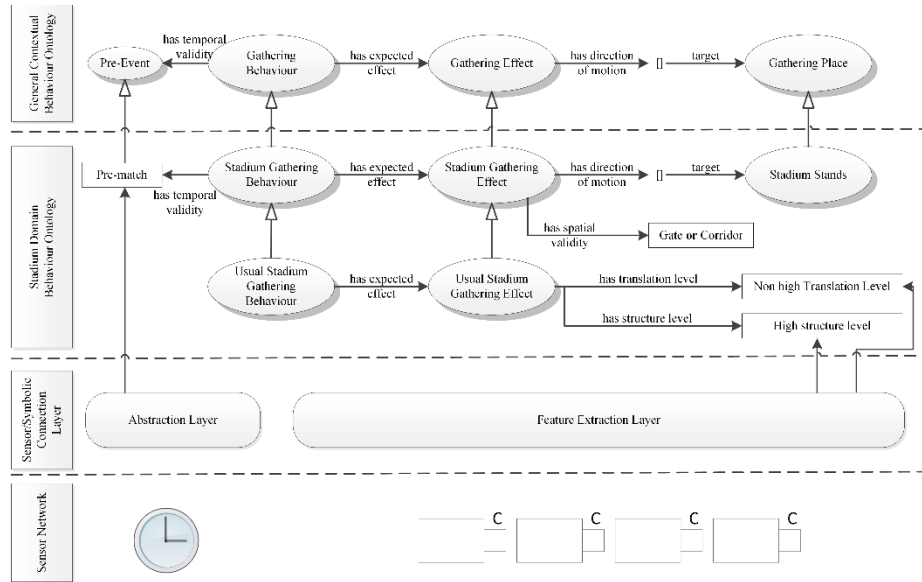

Fig. 2. Crowd Behaviour Ontology modules and Architecture

It is clear from **Fig. 2** that the described concepts of Gathering Behaviour, Effect, and Gathering Place in a general ontology can be imported by various domain ontologies that will refine their definition for specific venue types. Equally, the mentioned sensory video feeds and further context relevant information, such as time and event occurrences, need to go through abstraction steps with symbolic representations in the terms of the domain ontology used (i.e. pre-match symbol instead of a date time).

5 Conclusions

This early research describes an approach for semantically enhancing the ingestion of sensor data which may describe the links to crowd behaviour, context and detectable sensor observations. The approach is inspired from script theory which suggests that humans follow scripts which are linked to past experience and observed situations. The paper presents an approach which models such scripts as causal chains in an OWL ontology language. The ontology can be reused to describe domain specific behaviour(s) at different venues. Moreover, a general schema is introduced on how sensor data is processed prior to behaviour classification using the ontology.

References

1. R. C. Schank and R. P. Abelson, *Scripts, plans, and knowledge*. Yale University, 1975.
2. Y.-H. Chang and H.-H. Yang, "Aviation occupant survival factors: An empirical study of the SQ006 accident," *Accid. Anal. Prev.*, vol. 42, no. 2, pp. 695 – 703, 2010.
3. J. Gonzàlez i Sabaté, "Human sequence evaluation: the key-frame approach," info:eu-repo/semantics/doctoralThesis, Universitat Autonoma de Barcelona, 2004.
4. C. Fernández and J. Gonzàlez, "Ontology for semantic integration in a cognitive surveillance system," in *Semantic Multimedia*, Springer, 2007, pp. 260–263.
5. A. K. Dey and G. D. Abowd, "The context toolkit: Aiding the development of context-aware applications," in *Workshop on Software Engineering for wearable and pervasive computing*, 2000, pp. 431–441.
6. P. Remagnino, H. Hagras, S. Velastin, and N. Monekosso, "Ambient intelligence: a gentle introduction," 2005.
7. H. Chen, T. Finin, and A. Joshi, "An ontology for context-aware pervasive computing environments," *Knowl. Eng. Rev.*, vol. 18, no. 03, pp. 197–207, 2003.
8. T. Gu, H. K. Pung, and D. Q. Zhang, "A middleware for building context-aware mobile services," in *Vehicular Technology Conference, 2004. VTC 2004-Spring. 2004 IEEE 59th*, 2004, vol. 5, pp. 2656–2660.
9. A. ten Teije and F. van Harmelen, "An extended spectrum of logical definitions for diagnostic systems," in *Proceedings of DX-94 Fifth International Workshop on Principles of Diagnosis*, 1994, pp. 334–342.
10. S. Cox, "Observations and Measurements," Open GIS Consortium Inc., 03-022r3, 2003.
11. M. Compton, P. Barnaghi, L. Bermudez, R. García-Castro, O. Corcho, S. Cox, J. Graybeal, M. Hauswirth, C. Henson, A. Herzog, V. Huang, K. Janowicz, W. D. Kelsey, D. Le Phuoc, L. Lefort, M. Leggieri, H. Neuhaus, A. Nikolov, K. Page, A. Passant, A. Sheth, and K.

Taylor, "The SSN ontology of the W3C semantic sensor network incubator group," *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 17, pp. 25–32, Dec. 2012.

12. J. J. Fruin, "Designing for Pedestrians: A Level-of-Service Concept. New York Metropolitan Association of Urban Designers and Environmental Planners," *Highw. Res. Rec.*, no. 355, 1971.