FORMAL REPRESENTATION OF EVENTS IN A SURVEILLANCE DOMAIN ONTOLOGY

F. Sobhani, K. Chandramouli, Q. Zhang, E. Izquierdo

Multimedia and Vision Research Group,
School of Electronic Engineering and Computer Science,
Queen Mary, University of London,
Mile End Road, E1 4NS, London, UK

ABSTRACT

Following the exponential deployment of surveillance systems across a wide-spread region of geographic locations, detection and representation of events has become a critical element in automated surveillance systems. In this paper, we present an extensive ontology framework for representing complex semantic events. The proposed ontology builds on DOLCE ontology and relies on the linguistic and cognitive modelling of philosophical knowledge to achieve interoperability between proprietary surveillance systems. The explicit definition of event vocabulary presented in the paper is aimed at aiding forensic analysts to objectively identify and represent complex events. The expressiveness of the proposed ontology framework is described in the context of London Riots which took place in 2011.

Index Terms— Surveillance ontology, Forensic analyst, Event Detection, DOLCE

1. INTRODUCTION

In a world of heightened vandalism and terrorist activities, video surveillance forms an integral part of any incident investigation. Based on David Davis estimation, which is still unrivalled, the figure of how many CCTV cameras there are in London stands at around one camera for every 14 people, adding up to a total of about 422,000[1]. The British Security Industry Authority (BSIA) estimated there are up to 5.9 million closed-circuit television cameras in the country, including 750,000 in "sensitive locations" such as schools, hospitals and care homes[2]. Following the ever increasing deployment of CCTV, there is a critical need for developing an "automated video surveillance system" with the capability of detecting complex events to aid the forensic investigators in solving the criminal cases. As an example,

in the aftermath of the London riots in August 2011 police had to scour through more than 200,000 hours of CCTV to identify suspects. Around 5,000 offenders were found by trawling through the footage, after a process that took more than five months. Similarly, finding missing people is similarly arduous work - when teenager Alice Gross went missing in September last year 30 officers were tasked with combing through CCTV from 30 cameras, covering a six-mile radius[3]. As surveillance systems grow in scale, heterogeneity and utility, there is an increasingly critical need to provide automated and smart surveillance solutions. In an effort to develop an open and expandable video analysis framework equipped with tools for analysing, recognising, extracting and classifying events in video, which can be used for searching during investigations with unpredictable characteristics, or exploring normative (or abnormal) behaviours, several efforts for standardising event representation from surveillance footage have been made by the researchers. The approaches presented in the literature can be broadly classified into mark-up language based [1, 2, 3, 4, 5, 6] and Event model based [7, 8]. While several approaches has relied on offering foundational support for the domain ontology extension, to the best of our knowledge a systematic ontology framework for standardising the event vocabulary for forensic analysts has not been presented in the literature.

Addressing the critical need for the standardising the event vocabulary, in this paper, we present an ontology framework for the semantic retrieval of complex events. The ontology framework is a derivative of DOLCE foundational ontology aimed to represent events that forensic analysts commonly encounter to aid in the investigation of criminal activities. The systematic categorisation of vast number of events aligned with the philosophical and linguistic theories enables the ontology framework for interoperability between surveillance systems.

The remainder of the paper is organised as follows. The literature review presented in Section 2, Section 3 presents

---

[1]http://www.cctv.co.uk/how-many-cctv-cameras-are-there-in-london/

[2]http://www.telegraph.co.uk/technology/10172298/One-surveillance-camera-for-every-11-people-in-B ritain-says-CCTV-survey.html

---

[3]http://www.wired.co.uk/news/archive/2015-08/17/one-nation-under-cctv

a detailed description on the role of foundational ontology, the description of the rationale behind engineering methodology adopted for extending DOLCE upper ontology for event categorisation in surveillance domain is presented in section 4, Section 5 presents the validation of the proposed ontology based on OntoClean process along with functional evaluation of the ontology based on CCTV footage of London riots from 2011 and subsequently in Section 6 the conclusions and roadmap for future research is presented.

2. LITERATURE REVIEW

In [1], authors presented the Event Recognition Language (ERL), which can describe hierarchical representation of complex spatio-temporal and logical events. The proposed event structure consists of such units as primitive, single-thread, and multithread events. Another event representation ontology, called CASEE, is based on natural language representation proposed in [2] and then extended in [3]. Subsequently, in [4] Video Event Representation Language (VERL) was proposed for describing ontology of events and the complementary Video Event Markup Language (VEML), which is a representation language for describing events in video sequences based on OWL [5]. In [6], the event detection is performed using a set of rules in SWRL language.

The Event Model E [7] has been developed based on an analysis and abstraction of events in various domains such as research publications [9], personal media [10], meetings [11], enterprise collaboration [12], and sports [13]. The Ontology framework provides a generic structure for the definition of events and is extensible to the requirements of events in the most different concrete applications and domains. In order to address the shortcomings of the existing ontology frameworks to represent Events, in [8] authors present a formal model of events, called Event-Model-F. The model is based on the foundational ontology DOLCE+DnS Ultralite (DUL) and provides comprehensive support to represent time and space, objects and persons, as well as mereological, causal, and correlative relationships between events. In addition, the Event-Model-F provides a flexible means for event composition, modelling event causality and event correlation, and representing different interpretations of the same event. The Event-Model-F is developed following the pattern-oriented approach of DUL, is modularised in different ontologies, and can be easily extended by domain specific ontologies. The reported approaches in the literature provide a framework for the representation of event ontology. However, to the best of our knowledge, no standard formalisation of forensic events has been proposed. Therefore, the proposed complex event vocabulary categorisation of the forensic complex events, will aid in enabling interoperability between surveillance systems and offer semantic retrieval of respective events.

3. THE ROLE OF FOUNDATION ONTOLOGY

As a definition, ontology offers "formalisation of conceptualisation". However due to the over growing interest in researches to represent information using ontology, there is a wide gap generated between the agreement of concept formalisation. While such an impact is minimal in general ontologies, their effect are magnified while defining a domain ontology. For example, the concept DamageProperty and Vandalism often lead to mis-conception with respect to the subjectivity of the forensic analyst. From the ontology engineering perspective, DamageProperty is a subclass of Vandalism, therefore every instance of DamageProperty is also an instance of Vandalism. However, the inverse need not be held true, as Vandalism can also represent breaking a window. Therefore, in order to eliminate the terminological ambiguity the domain ontology concepts need to be aligned with the terminology of foundational ontology. Such a process will facilitate mutual understanding and inter-interoperability among people and machines [14]. In this regard, several efforts have been taken by researches in defining the foundational ontologies, such as BFO[4], SUMO[5], UFO[6], DOLCE[7], to name a few. Since DOLCE ontology offers a cognitive bias with the ontological categories underlying natural language and human common sense, the same is selected for our proposed extension. As it is shown in Fig.1, *DOLCE* foundational ontology encompasses *Endurant* and *perdurant* entities. *Endurant* entities are ever-present at any time as opposed to *perdurant* entities, which are known as processes, events, states and activities extended in time by accumulating different temporal parts. A more thorough explanation on the DOLCE conceptualisation can be found in [15].

4. FORENSIC COMPLEX EVENT ONTOLOGY

In order to achieve the formalisation of complex event categorisation, the proposed vocabulary structure of human actions extends on the *perdurant* entities of DOLCE ontology.

Vendler [16][17] described a 4-way classification of action-verbs into *states*, *activities*, *achievements* and *accomplishment*. These recommendations as presented in [18] are followed to assign the action classes into respective categories. The classification is achieved based on the representation of event properties namely, telic, stage and cumulative. The fundamental characteristic distinction between these concepts are derived from the event properties and are listed in Table 1.
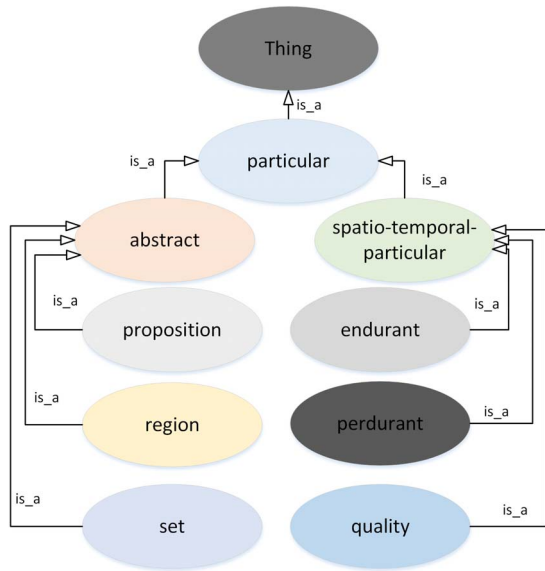
- **State[-telic,-stages]** The action category represents a long, non-dynamic event in which every instance is the

---
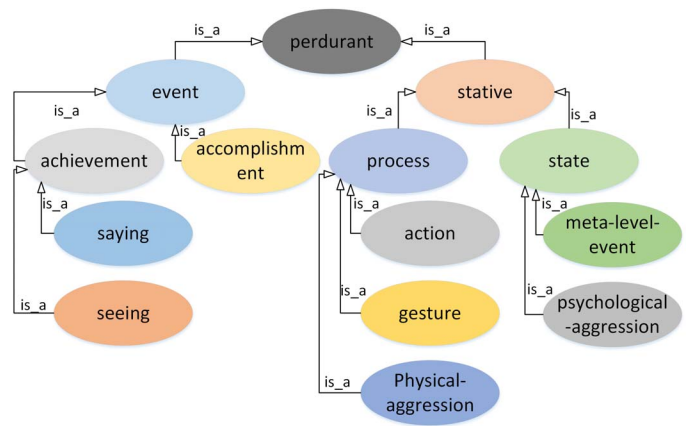
Fig. 1. The high-level DOLCE Ontology concept overview

Fig. 2. The concept description of Perdurant for forensic event decomposition

Table 1. Classification of Event Types

State	-Telic	-Stage	Cumulative
Activity	-Telic	+Stage	-
Achievement	+Telic	-Stage	Not Cumulative
Accomplishment	+Telic	+Stage	Not Cumulative

same representation as every other bit and there can not be any distinction made between the stages. States are cumulative and homogenous in nature.

- **Process [-telic, +stages]** The action category, like state, is atelic, but unlike state, the action undertaken are dynamic. The actions appear progressively and thus can be split into a set of stages for analysis.

- **Accomplishments [+telic, +stages]** Accomplishments are telic and cumulative activity, and thus behave differently from both State and Process. The performed action can be analysed in stages and in this way they are similar to Process. Intuitively, an accomplishment is an activity which moves toward a finishing point as it has variously been called in the literature. Accomplishment is a also cumulative activity.

- **Achievements [+telic, stages]** Achievements are similar to accomplishments in their telicity. They are also not cumulative with respect to contiguous events achievements do not go on or progress, because they are near instantaneous, and are over as soon as they have begun.

The forensic ontology proposed in this paper strictly adheres to the above terminological determination of action

categories as mentioned above and extends the classes with suitable event concepts. As it is presented in Fig.2, the concept "State" offers representation for "MetaLevelEvent" which encompasses abstract human events such as "Accusing", "Believing" and "Liking" among others. As previously stated State represents a collection of events which are exhibited by human that are time-consuming, non-dynamic, cumulative and homogenous. The other sub-class of "State" is "PsychologicalAggression" which characterises the human actions such as "Blaming", "Decrying", "Harassing" and so forth.

The concept "Process" includes several human action categories that represent dynamic events which can be split into several intermediate stages for analysis. For the purposes of clarity, the concept process offers three sub-concepts namely "Action", "Gesture" and "PhysicalAggression". The "Action" class incorporates different event such as "Dancing", "Greeting", "Hugging" among other concepts defined. The concept "Gesture" formalises the different interest points related to human gestures. In order to eliminate the ambiguity traditionally present in human gestures across cross-cultural impact, the action performed during the gesture is captured and represented in the Ontology. Thus enabling the removal of subjectivity from the concept definition. The final sub-class of process includes "PhysicalAggression" and thus formalises human actions resulting in conflicts.

By and large, the human action categorised into State and Process represent the microscopic movements of humans. From the automatic surveillance viewpoint, these microscopic events are extracted from the media items. In contrast, the event representation formalised among "Achievement" and "Accomplishment" offer a rich combination of human events that allow for the construction of complex events with or without the combination of microscopic features. The concept hierarchy for Vandalism, as a subclass of "Accom-

Fig. 4. The functional evaluation of the proposed ontology with a consideration of a use-case from the London Riots 2011
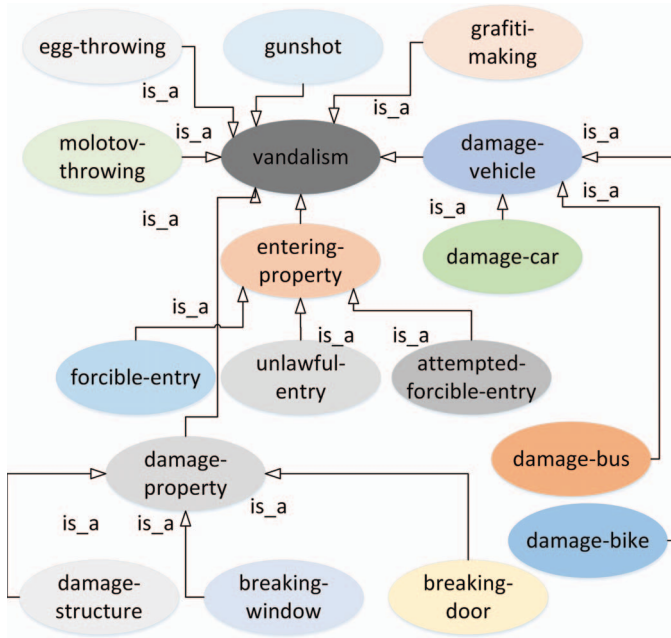

Fig. 3. The concept hierarchy for Vandalism, which is a sub-class of "Accomplishment"

plishment" is illustrated in Fig.3.

5. ONTOLOGY VALIDATION

For the intention of achieving inter-operability between system, it is important to evaluate the proposed ontology model using triple synergy namely structural, functional and usability-profile as suggested in [19]. The structural dimension of ontologies focuses on syntax and formal semantics and in this form the topological, logical and meta-logical properties of an ontology can be measured by means of context-free metric. The functional dimension is related to the intended use of a given ontology and of its components. Finally, the usability dimension focusses on the ontology profile which typically addresses the communication context of an ontology. For the structural evaluation of the proposed ontology, the axiomatic suggestions of OntoClean [20], methodology is used and the properties are tagged with the meta-properties namely Rigidity, Identity, Unity and Dependence. The functional evaluation of the proposed ontology is presented with a consideration of a use-case (depicted in Fig. 4) from the London Riots 2011 with the mapping of visual events to the concepts proposed in the Ontology[8].

6. CONCLUSION AND FUTURE WORK

In this work, we have proposed the extensive ontology framework for representing complex semantic event. The proposed ontology builds on DOLCE foundational ontology and relies on the linguistic and cognitive modelling of philosophical knowledge. The proposed ontology structure aims at formalising the events which are often required by forensic analysts to identify from a crime scene to enable them track the person of interest. In this context, the perdurant as defined in the DOLCE ontology as "occurents in time" are extended to represent events categorised into static, events, accomplishments and achievements. As aligned with the definition of the concepts, the representative events from the CCTV footage are instantiated against the corresponding concepts from the proposed ontology. Such a functional representation eliminates the disambiguation of concept definition and allows for the inter-operability of automated surveillance system. Additionally, the ontology has been validated using a use-case from the London Riot 2011 and can be used for the similar use-cases.

---

[8]The complete ontology model offers the forensic analysts an opportunity to express complex events. The event model of the ontology is attached as supplementary material. The ontology model will be published upon the acceptance of the paper.

7. REFERENCES

[1] R. Nevatia, T. Zhao, and S. Hongeng, "Hierarchical language-based representation of events in video streams," 2003.

[2] A. Hakeem, Y. Sheikh, and M. Shah, "Casee: A hierarchical event representation for the analysis of videos," in *in: The Nineteenth National Conference on Artificial Intelligence*, 2004.

[3] A. Hakeem, K. Shafique, and M. Shah, "An object-based video coding framework for video sequences obtained from static cameras," in *Proceedings of the 13th Annual ACM International Conference on Multimedia*, ser. MULTIMEDIA '05. New York, NY, USA: ACM, 2005, pp. 608–617. [Online]. Available: http://doi.acm.org/10.1145/1101149.1101289

[4] R. Nevatia, J. Hobbs, and B. Bolles, "An ontology for video event representation," in *Computer Vision and Pattern Recognition*. IEEE, 2004.

[5] A. R. J. Francois, R. Nevatia, J. Hobbs, and R. C. Bolles, "Verl: An ontology framework for representing and annotating video events," *IEEE MultiMedia*, vol. 12, no. 4, pp. 76–86, Oct. 2005. [Online]. Available: http://dx.doi.org/10.1109/MMUL.2005.87

[6] L. Snidaro, M. Belluz, and G. L. Foresti, "Representing and recognizing complex events in surveillance applications," *2014 11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, vol. 0, pp. 493–498, 2007.

[7] U. Westermann and R. Jain, "Toward a common event model for multimedia applications," *MultiMedia, IEEE*, vol. 14, no. 1, pp. 19–29, Jan 2007.

[8] A. Scherp, T. Franz, C. Saathoff, and S. Staab, "F–a model of events based on the foundational ontology dolce+dns ultralight," in *Proceedings of the Fifth International Conference on Knowledge Capture*, ser. K-CAP '09. New York, NY, USA: ACM, 2009, pp. 137–144. [Online]. Available: http://doi.acm.org/10.1145/1597735.1597760

[9] B. Gong, R. Singh, and R. Jain, "Researchexplorer: Gaining insights through exploration," in *in Multimedia Scientific Data, Proc. 6th ACM SIGMM International Workshop on Multimedia Information Retrieval*, p. 7.

[10] P. Appan and H. Sundaram, "Networked multimedia event exploration," in *Proceedings of the 12th Annual ACM International Conference on Multimedia*, ser. MULTIMEDIA '04. New York, NY, USA: ACM, 2004, pp. 40–47. [Online]. Available: http://doi.acm.org/10.1145/1027527.1027536

[11] R. Jain, P. Kim, and Z. Li, "Experiential meeting system," in *Proceedings of the 2003 ACM SIGMM Workshop on Experiential Telepresence*, ser. ETP '03. New York, NY, USA: ACM, 2003, pp. 1–12. [Online]. Available: http://doi.acm.org/10.1145/982484.982486

[12] P. Kim, M. Podlaseck, and G. Pingali, "Personal chronicling tools for enhancing information archival and collaboration in enterprises," in *Proceedings of the the 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences*, ser. CARPE'04. New York, NY, USA: ACM, 2004, pp. 56–65. [Online]. Available: http://doi.acm.org/10.1145/1026653.1026662

[13] G. Pingali, A. Opalach, Y. Jean, and I. Carlbom, "Instantly indexed multimedia databases of real world events," *Multimedia, IEEE Transactions on*, vol. 4, no. 2, pp. 269–282, Jun 2002.

[14] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari, "Wonderweb deliverable d18, ontology library (final)," *ICT project*, vol. 33052, 2003.

[15] R. Casati and A. Varzi, "Events," in *The Stanford Encyclopedia of Philosophy*, winter 2015 ed., E. N. Zalta, Ed., 2015.

[16] Z. Vendler, "Verbs and times," *The philosophical review*, pp. 143–160, 1957.

[17] Z. Vendler, Ed., *Linguistics in philosophy*, ser. G - Reference, Information and Interdisciplinary Subjects Series. Cornell University Press, 1967. [Online]. Available: https://books.google.co.uk/books?id=OR1DAAAAIAAJ

[18] S. Rothstein, "Verb classes and aspectual classification," *Structuring Events: A Study in the Semantics of Lexical Aspect*, pp. 1–35, 2004.

[19] A. Gangemi, C. Catenacci, M. Ciaramita, and J. Lehmann, "Modelling ontology evaluation and validation," in *The Semantic Web: Research and Applications*, ser. Lecture Notes in Computer Science, Y. Sure and J. Domingue, Eds. Springer Berlin Heidelberg, 2006, vol. 4011, pp. 140–154.

[20] N. Guarino and C. Welty, "Evaluating ontological decisions with ontoclean," *Commun. ACM*, vol. 45, no. 2, pp. 61–65, Feb. 2002. [Online]. Available: http://doi.acm.org/10.1145/503124.503150