Goal, Video Description and Capability Ontologies for Fish4Knowledge Domain*

Gayathri Nadarajan, Yun-Heh Chen-Burger

School of Informatics, University of Edinburgh, U.K.
gaya.n@ed.ac.uk, jessicac@inf.ed.ac.uk

Abstract. We created a set of domain ontologies that are based on user requirements for the Fish4Knowledge (F4K) project – goal, video description and capability. The roles of the ontologies are to 1) support the development of appropriate functions of the project's workflow system, and 2) serve as a communication media to interface with other F4K components. The ontologies were designed with collaboration with image processing experts, marine biologists and user interface experts to capture the domain knowledge succinctly. They were utilised in the first version of our workflow composition and execution system for video classification, fish detection and counting tasks. They will continue to evolve with F4K's needs and are envisaged to interface with other components.

Keywords: ontologies, semantics based workflows, requirements based virtual workflow system, intelligent video processing.

1 Introduction

The workflow component of the Fish4Knowledge (F4K) project is responsible for the composition and execution of a set of video and image processing (VIP) modules on high performance computing machines based on user requirements and descriptions of the video data. The workflow component interprets the user requirements as high level VIP tasks, create workflows based on the procedural constraints of the VIP modules to ultimately invoke and manage their execution in a distributed environment. In our intelligent workflow system, we have chosen to use an ontological-based approach to guide the automatic generation of a "virtual workflow machine" based on a set of closely related ontologies. This allows a separation between the problem and application descriptions and the workflow mechanism. As a result, the workflow machine may work in different problem domains if the problem and application descriptions are changed. Consequently, this will promote reusability and provide a conceptualisation that can be used between different domain experts, such as marine biologists, image processing experts and workflow engineers. These ontologies are also pivotal for reasoning. For instance, in the selection of optimal VIP software modules, the Capability Ontology is used to record known heuristics obtained from VIP experts.

* This research was funded by European Commission FP7 grant 257024, in the Fish4Knowledge project (www.fish4knowledge.eu).

This paper describes our efforts in creating a set of suitable domain ontologies that are based on user requirements (from marine biologists) for our intelligent workflow system, the ontologies will 1) support the development of appropriate functions of the workflow system, and 2) serve as a communication media to interface with other F4K components. We first describe the 20 scientific questions provided by F4K's marine biologists, and then provide an analysis of these questions from the workflow system's point of view – that is based on the capabilities of the VIP modules available to us (Section 2). Based on a mapping between the user requirements and a high level abstraction of the capabilities of the VIP modules, we have constructed the **Goal Ontology** (Section 3). The **Video Description Ontology** (Section 4) contains the environmental factors related to the videos, among others. The **Capability Ontology** (Section 5) describes details of the VIP tools and techniques. For ontology development and visualisation purposes, OWL 1.0 [3] was generated using Protege version 4.0. Where applicable, ontological diagrams were derived using the OntoViz plugin.

To date, the Goal Ontology contains 52 classes, 85 instances and 1 property, the Video Description Ontology has 24 classes, 30 instances and 4 properties and the Capability Ontology has been populated with 42 classes, 71 instances and 2 properties. They have supported the first version of the workflow system that has been evaluated for efficiency, adaptability and user learnability in video classification, fish detection and counting tasks [4, 5].

2 Related Work

Ontologies have been used for describing VIP domain knowledge by several major projects, such as aceMedia [1], VIDI-Video [2], and NeOn [6], just to name a few. VIDI-Video and aceMedia have worked on developing multimedia ontologies for video annotation tasks. The aceMedia project has been a successful effort in utilising ontologies to capture high level descriptions to guide the low level technical tasks of multimedia annotation via their core, domain and multimedia ontologies, all described in RDFS. The VIDI-Video project uses semantics-based and machine learning technologies to explore the area of semantics-based video search. The NeOn project has developed a network of fisheries ontologies that has some relevance to our work. We have identified and intergrated FAO's concepts and properties that were relevant to F4K. These include fishing areas[1], land areas[2] and biological entities[3]. This effort is the first step in integrating our work with existing standards in the ontological field.

3 Goal Ontology

The Goal Ontology contains the high level questions posed by the user, interpreted by the system as VIP tasks (termed as goals) and the constraints to the goals. Before it could be constructed, a list of queries posed by marine biologists were examined. These questions are listed in Table 1.

1 http://www.fao.org/fishery/cwp/handbook/H/en
2 http://www.fao.org/countryprofiles/geoinfo.asp
3 http://www.fao.org/fishery/collection/asfis/en

Table 1. List of 20 scientific questions posed by F4K's Taiwanese marine biologists.

Q1	How many species appears and their abundance and body size in day and night including sunrise and sunset period.
Q2	How many species appears and their abundance and body size in certain period of time (day, week, month, season or year). Species composition change within one period.
Q3	Give the rank of above species, *i.e.* listed according to their abundance or dominance. How many percent are dominant (abundant), common, occasional and rare species.
Q4	Fish color pattern change and their behavior in the night for diurnal fish and vice versa for nocturnal fishes.
Q5	Fish activity within one day (24 hours).
Q6	Feeding, predator-prey, territorial, reproduction (mating, spawning or nursing) or other social or interaction behavior of various species.
Q7	Growth rate of certain species for a certain colony or group of observed fishes.
Q8	Population size change for certain species with one period of time.
Q9	The relationship of above population size change or species composition change with environmental factors, such as turbidity, current velocity, water temperature, salinity, typhoon, surge or wave, pollution or other human impact or disturbance *etc.*
Q10	Immigration or emigration rate of one group of fishes inside one monitoring station or one coral head.
Q11	Solitary, pairing or schooling behavior of fishes.
Q12	Settle down time or recruitment season, body size and abundance for various fishes.
Q13	In certain area or geographical region, how many species could be identified or recognized easily and how many species are difficult. The most important diagnostic character to distinguish some similar or sibling species.
Q14	Association among different fish species or fish-invertebrates.
Q15	Short term, mid-term or long term fish assemblage fluctuation at one monitoring station or comparison between experimental and control (MPA) station.
Q16	Comparison of the different study result between using diving observation or underwater real time video monitoring techniques. Or the advantage and disadvantage of using this new technique.
Q17	The difference of using different camera lens and their angle width.
Q18	Is it possible to do the same monitoring in the evening time.
Q19	How to clean the lens and solve the biofouling problem.
Q20	Hardware and information technique problem and the possible improvement based on current technology development and how much cost they are.

As a starting point, the 20 questions provided by the marine biologists were mapped to the VIP tasks (goals) that they fall into based on the types of tasks that are required to be carried out to address these questions. A VIP task as indicated here is a high level goal in our Goal Ontology that can be understood by an image processing expert, who writes low level programs or software to solve it. Table 2 contains the result of the analysis after communication with image processing experts, marine biologists and user interface experts. The queries are indicated as Q1 to Q20 and the eight main VIP goals are contained in columns 2 to 9 of the table.

Table 2: Mapping of high-level VIP tasks to the 20 questions as provided by Fish4Knowledge's marine biologists.

Query	Fish Detection and Tracking	Population Related Analysis	Fish Species Classification	Behaviour Understanding	Fish Clustering	Fish Feature Analysis	Event Detection	Aggregated Analysis	Assumptions/Definitions
1	x	x	x			x			Assumption: Sunrise & sunset are defined by IP teams as a 10-minute period when videos change from dark to colour & vice-versa. Night videos are not available.
2	x	x	x					x	Definition: *Species composition* is defined as the number of fish from each species.
3	x	x	x	(x)		x		x	
4	x	x	x	x				x	Comment: *Diurnal* and *nocturnal* fishes will need to be determined manually by F4K researchers and populated in the database. Colour pattern change may not be possible to be performed as the change will occur within 1–2 seconds, which is too quick for current VIP techniques to detect.
5	x	x	x	x		x			Assumption: *Fish activity* refers to activity of *all* the fish in a video clip.
6	x	x	x	x		x		(x)	Comment: Territorial behaviour is seen when a fish species attack other fish or human entering the area that they are in.
7	x	x	x			x		x	Definition: *Fish group* is a group of fish (of the same or closely related species) living together in one area. Definition: *Fish colony* refers to the same fish species. Definition: Growth rate = birth rate – death rate + immigration rate – emigration rate
8	x	x	x		(x)	x	(x)	x	Assumption: *Population size change* depends on (not exclusively, but most commonly) immigration, emigration, birth rate and death rate.
9	x	x	x	(x)		x	x	x	Comment: NCHC could provide water salinity, temperature, turbidity, ph value and dissoved oxygen. In the future typhoon incidence is a possibility.
10	x	x	(x)			(x)		x	Comment: Mathematical formula for *immigration* and *emigration rate* are required for aggregated analysis.
11	x	x	x	x		x			
12	x	x	x			x		x	Definition: *Settle down time* refers to the time it takes to colonise a new area through immigration. In marine terms, settlement refers to the colonisation of a previously unavailable area, *e.g.* new structure

Table 2 – Continued

Query	Fish Detection and Tracking	Population Related Analysis	Fish Species Classification	Behaviour Understanding	Fish Clustering	Fish Feature Analysis	Event Detection	Aggregated Analysis	Assumptions/Definitions
									(sunken ship) or an area that has been cleared (usually by a destructive event, *e.g.* storms, human action). Definition: *Recruitment season* is the phenomenon where new born fishes survive and are added to the overall population.
13	x	x	x		x	x		x	Comment: Similar/sibling species could be physically similar (*e.g.* shape, size, number of spikes), or having the same genus but not species. Comment: This task requires labelled data, *i.e.* ground truth in order to determine whether a fish species is easy or difficult to be recognised.
14	x	x	x	x	(x)	x		(x)	Definition: *Association* refers to the co-occurrence of species. It could be considered with the number of co-occurrences per associated species, average time-span and threshold of time-span between two occurrences. An example is the symbiotic relationship between fish and fish-invertebrates.
15	x	x	x	(x)		x		x	Definition: A *fish assemblage* is a group of interacting populations within a specific area. They must interact *e.g.* predator/prey interactions, compete for food, space or other resources. Fish assemblage fluctuation is the change within the structure of the assemblage, *e.g.* increase/decrease in numbers either either overall or within individual populations.
16								x	Comment: This task requires results from manual processing in order for comparisons to be performed.
17								x	Comment: This is not possible at present.
18								x	Comment: This is not possible at present unless infra-red cameras are installed and recordings are provided.
19							x	x	Comment: This question should be rephrased as 'When to clean the lens'. How to clean the lens and solve the bio-fouling problem is beyond F4K's IP capabilites.
20								x	Comment: This task is not possible at present.

For each query the relevant VIP task(s) that can be used to support it are marked with an 'x'. An 'x' marked in parentheses indicate that the particular VIP task could be associated with the query, but it could not be determined with absolute certainty that this function will be supported within this project.

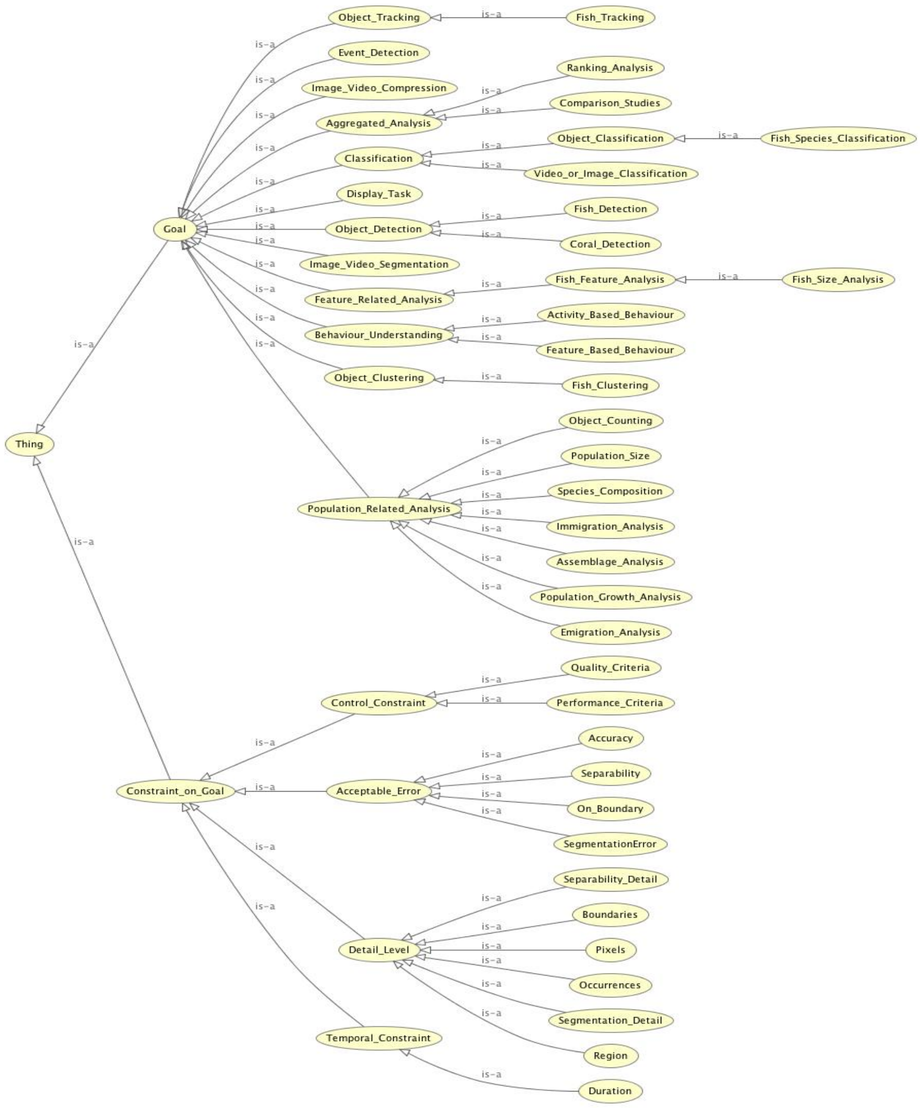

Fig. 1. Goal ontology denoting the main classes of goals and constraints that were applicable to the 20 questions posed by Taiwanese marine biologists.

This may be due to technical reasons or the time constraints imposed on this project. The VIP tasks identified here constitute the main goals that we describe and use in the Goal Ontology. Along with these main goals, we add higher level goals to provide a classification structure and linkage to possible user queries. Fig. 1 depicts the Goal Ontology pictorially. The main relationship between classes shown here is the sub-class relationship, *is-a*. Instances and properties are not shown due to the limitation of the visualisation tool.

The 'Goal' class is the umbrella of concept that includes the VIP tasks identified in Table 2. The eight main VIP tasks can be found in Figure 1. We also created a level of intermediate classes between the goal class and the lower level VIP goals to increase the flexibility and readability of the ontology. Under these general concepts, more specific goals may be defined, for example 'Fish Detection', 'Fish Tracking', 'Fish Clustering', 'Fish Species Classification' and 'Fish Size Analysis'. The principle behind keeping the top level concepts more general is also to allow the ontology to be easily extended to include other (new) tasks as appropriate as the project develops.

'Constraint on Goal' refers to the conditions that restrict the video and image processing tasks or goals further. In F4K's context, the main constraint for a VIP goal is the 'Duration', a subclass of 'Temporal Constraint'. Each task may be performed on all the historical videos, or a portion specified by the user – within a day, night, week, month, year, season, sunrise or sunset (all specified as instances of the class 'Duration'). Other constraints types include 'Control Constraint', 'Acceptable Error' and 'Detail Level'. The control constraints are those related to the speed of VIP processing and the quality of the results expected by the user. Acceptable errors are the threshold for errors that the user may want to insert. An example of this is the criterion 'Accuracy', which states the accuracy level of a detected object. The class 'Detail Level' contains constraints that are specific to particular details, for example detail of 'Occurrence' is used for detection tasks to constrict the number of objects to be detected.

The Goal Ontology is used for consistency checks when a user query is detected in the system. It can check that the query matches with a goal or set of goals that is achievable within the workflow system. It is also used to guide the selection of higher level tasks for workflow and formulate input values to the reasoning engine that is responsible for searching the VIP solution set for a VIP task, *i.e.* to compose the workflow.

4 Video Description Ontology

The Video Description Ontology describes the concepts and relationships of the video and image data, such as what constitutes video/image data, the acquisition conditions such as lighting conditions, colour information, texture, **environmental conditions** as well as spatial relations and the range and type of their values. Fig. 2 gives a pictorial overview of the main components of the Video Description Ontology. The upper level classes include 'Video Description', 'Descriptor Value', 'Relation', and 'Measurement Unit'.

The main class of this ontology is the 'Video Description' class, which has two subclasses – 'Description Element' and 'Descriptor'. A description element can be either a 'Visual Primitive' or an 'Acquisition Effect'. A visual primitive describes visual effects of a video/image such as observed object's geometric and shape features, *e.g.* size, position and orientation while acquisition effect descriptor contains the non-visual effects of the whole video/image that contains the video/image class such as the brightness (luminosity), hue and noise conditions.

The descriptor for the description elements are contained under the 'Descrip-

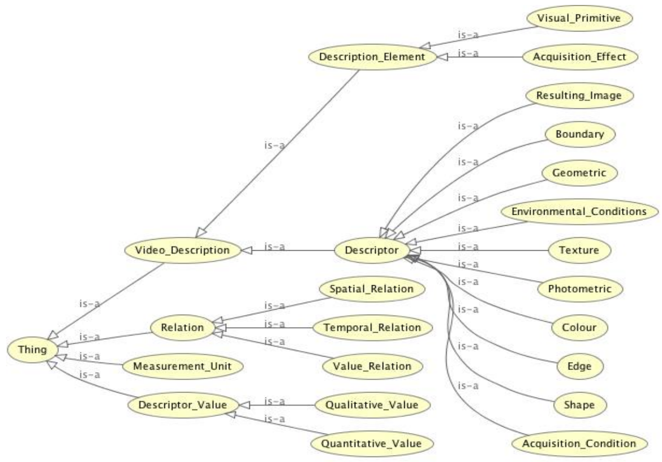

Fig. 2. Main concepts of the Video Description Ontology.

tor' class and are connected to the 'Description Element' class via the object property 'hasDescriptionElement'. Typical descriptors include shape, edge, colour, texture and environmental conditions. Environmental conditions, which are acquisitional effects, include factors such as current velocity, pollution level, water salinity, surge or wave, water turbidity, water temperature and typhoon.

These values that the descriptors can hold are specified in the 'Descriptor Value' class and connected by the object property 'hasValue'. For the most part, qualitative values such as 'low', 'medium' and 'high' are preferred to quantitative ones (*e.g.* numerical values). 'Qualitative' values could be transformed to quantitative values using the 'convertTo' relation. This would require the specific measurement unit derived from one of the classes under the concept 'Measurement Unit' and conversion function for the respective descriptor *e.g.* a low velocity could be interpreted as movement with velocity within a range of 0 and 25ms^{-1}[4]. Some descriptor values can be tied to their appropriate measurement units. The property that specifies this is 'hasMeasurementUnit', which relates instances in the class 'Descriptor' to instances in the class 'Measurement Unit'. This ontology can be used to describe the videos and external effects on it such as environmental conditions. As with the Goal Ontology it is used for consistency checking when a set of video descriptors are supplied to the system from the user interface.

5 Capability Ontology

The Capability Ontology (Fig. 3) contains the classes of video and image processing tools, techniques and performance measures of the tools with known

[4] Currently, there is no fixed conversion formula. The actual conversion formula used will be determined as we gain more experience using our workflow system over time.

Fig. 3. Main concepts of the Capability Ontology.

domain heuristics. This ontology will be used to identify the tools that will be used for workflow composition and execution of VIP tasks. In our context, it is used by a reasoner for the selection of optimal VIP tools. The main concepts intended for this ontology have been identified as 'VIP Tool', 'VIP Technique' and 'Domain Descriptions for VIP Tools'. Each VIP technique can be used in association with one or more VIP tools. A VIP tool is a software component that can perform a VIP task independently, or a function within an integrated vision library that may be invoked with given parameters. 'Domain Description for VIP Tool' represent a combination of known domain descriptions (video descriptions and/or constraints to goals) that are recommended for a subset of the tools. This will be used to indicate the suitability of a VIP tool when a given set of domain conditions hold at a certain point of execution. At present these domain descriptions are represented as strings and tied to VIP tools, *e.g.* Gaussian background model would have the description 'Clear and Fast Background Movement' to indicate the best domain scenario for it to be selected.

The main types of VIP tools are video analysis tools, image enhancement tools, clustering tools, image transform tools, basic structures and operations tools, object description tools, structural analysis tools and object recognition and classification tools. At present fish detection and tracking have been performed more than the other tasks within Fish4Knowledge. Hence the ontology has been populated with most of the tools associated with these tasks. For other tasks that have not been performed, *e.g.* fish clustering, the ontology will be extended and populated in due course. Detection and tracking tools fall under the class 'Video Analysis Tool'. Other types of video analysis tools are event detection tools, background modelling tools and motion estimation tools.

The class 'Object Description Tool' specifies tools that extract features such as colour, texture, size and contour, while image transform tools are those concerned with operations such as point, geometric and domain transformations. 'VIP Technique' is a class that contains technologies that can perform VIP operations. For now, two types of machine learning techniques have been identified. These techniques could be used to accomplish the task of one or more VIP tools. For example, neural netwworks can be used as classifiers as well as detectors.

The Capability Ontology can be used for reasoning during workflow composition using planning. As planning takes into account preconditions before selecting a step or tool, it will assess the domain conditions that hold to be used in conjunction with an appropirate VIP tool.

6 Conclusions

We have presented three ontologies that relate to the goals, capabilities and environmental conditions within Fish4Knowledge. Development of these ontologies act as a starting point to bring various expertise within the project together (*e.g.* image processing, workflow, marine science and user interface). At its current status, these ontologies have been used to guide the construction of VIP workflows for single videos on single CPU machines. New workflows will be handling multiple videos on a high performance computing environment, which would require appropriate prioritising, scheduling, monitoring and fault tolerance management strategies. At present the ontologies are interfacing the workflow and the image processing components, it is envisaged that the ontologies would also serve as a communication media to interface with other F4K components in due course.

References

1. aceMedia: Integrating Knowledge, Semantics and Content for User-Centred Intelligent Media Services (2004), `http://www.acemedia.org/aceMedia`.
2. Bertini, M., Bimbo, A.D., Serra, G., Torniai, C., Cucchiara, R., Grana, C., Vezzani, R.: Dynamic Pictorially Enriched Ontologies for Digital Video Libraries. IEEE MultiMedia 16(2), 42–51 (2009)
3. McGuinness, D., van Harmelen, F.: OWL Web Ontology Language. World Wide Web Consortium (W3C) (2004), `http://www.w3.org/TR/owl-features`
4. Nadarajan, G., Chen-Burger, Y.H., Fisher, R.B.: A Knowledge-Based Planner for Processing Unconstrained Underwater Videos. In: IJCAI'09 Workshop on Learning Structural Knowledge From Observations (STRUCK'09) (2009)
5. Nadarajan, G., Chen-Burger, Y.H., Fisher, R.B.: SWAV: Semantics-based Workflows for Automatic Video Analysis. In: Special Session on Intelligent Workflow, Cloud Computing and Systems, (KES-AMSTA'11) (2011)
6. NeOn: Networked Ontologies Project (2006-10), `http://www.neon-project.org`