Rules-By-Example – a Novel Approach to Semantic Indexing and Querying of Images

Suzanne Little[1,2], Jane Hunter[1]

[1] DSTC Pty Ltd, The University of Queensland,
St Lucia, Queensland, Australia
{slittle, jane}@dstc.edu.au
[2] ITEE Dept., The University of Queensland,
St Lucia, Queensland, Australia

Abstract: Images represent a key source of information in many domains and the ability to exploit them through their discovery, analysis and integration by services and agents on the Semantic Web is a challenging and significant problem. To date the semantic indexing of images has concentrated on applying machine-learning techniques to a set of manually-annotated images in order to automatically label images with keywords. In this paper we propose a new hybrid, user-assisted approach, Rules-By-Example (RBE), which is based on a combination of RuleML and Query-By-Example. Our RBE user interface enables domain-experts to graphically define domain-specific rules that can infer high-level semantic descriptions of images from combinations of low-level visual features (e.g., color, texture, shape, size of regions) which have been specified through examples. Using these rules, the system is able to analyze the visual features of any given image from this domain and generate semantically meaningful labels, using terms defined in the domain-specific ontology. We believe that this approach, in combination with traditional solutions, will enable faster, more flexible, cost-effective and accurate semantic indexing of images and hence maximize their potential for discovery, re-use, integration and processing by Semantic Web services, tools and agents.

1 Introduction

Images constitute a significant and under-utilized proportion of information in many domains and semantic indexing of images is essential if they are to be discovered and fully exploited by Web search engines, services, agents and applications. Because of the complexity and multidimensional nature of image data, manual annotation is slow, expensive and predisposed to high subjectivity. Significant progress has been made in recent years on the automatic recognition of low-level features within images. However, comparatively little progress has been made on the machine-generation of high-level semantic descriptions of images.

In this paper we describe a unique, user-assisted approach to generating ontology-based semantic descriptions of images from low-level automatically extracted features. Our approach enables domain-experts to define rules specific to their domain which map particular combinations of low-level visual features (colour, texture, shape, size) to high-level semantic terms defined in their domain ontology. Such descriptions enable more sophisticated semantic querying of the images in terms familiar to the user's domain whilst still ensuring that the information and knowledge within the images has a much greater chance of being discovered and exploited by services, agents and applications on the Web. The use of ontologies reduces the potential subjectivity of the semantic descriptions and the documentation of the inference rules provides a mechanism for capturing and sharing human domain knowledge, in the form of the analysis methods employed by the domain experts.

Query-By-Example (QBE) is a well-recognized approach to image retrieval. In QBE, the user provides the system with an example image or examples of the visual features that they are seeking and the system attempts to retrieve similar images by matching the visual features. In order to define semantic inferencing rules for images, users need to be able to specify values for low-level visual features. Consider, for example, an oncologist labelling CAT scans of brains, in order to enable the search and retrieval of particular types of brain tumours:

IF [(color is like this) AND (texture is like this) AND (shape is like this)]
THEN (the object is an astrocytoma)

The simplest and most intuitive way to specify such rules is to provide a QBE-type interface by which users can specify colour, texture, shape and size through visual examples, drawing from images within their collection of interest. Hence we decided to develop a Graphical User Interface (GUI) which provides users with color palettes, texture palettes, pre-defined shapes or drawing tools, by which they can define their own semantic inferencing rules. The GUI is generated from and dependent on the pre-selected back-end (OWL) ontology (specified at time of system configuration) which defines and constrains the semantic descriptions that can be generated. The system can be migrated to a different domain by supplying a different back-end ontology. Graphically-defined rules are saved as RuleML which can be applied to image metadata from that particular domain to generate semantic descriptions.

In order to test and evaluate our proposed approach we wanted a domain in which image analysis played a key role and in which a certain amount of knowledge was already available in the form of existing but informal models, vocabularies and image analysis methods which could be represented within ontologies and inference rules. Discussions with the Centre for Microscopy and Microanalysis at The University of Queensland led us to the fuel cell community and their problem of optimization of fuel cells. This application is typical of many eScience applications in that it requires the analysis of large numbers of microstructural images and their assimilation with related information such as performance data and manufacturing parameters. Because the efficiency of fuel cells is dependent on the internal structure of the fuel cell layers and the interfaces between them, an analysis of electron microscopy images of cross-sectional samples through fuel cells can reveal valuable new information. Simple macro-level information such as the thickness of the cell layers, surface area, roughness and densities can be used to determine gas permeation of the electrode materials. Nano-level information about the electrode's internal interface structure provides data on the efficiency of exchange reactions. Figure 1 illustrates a sample of the image data obtainable at different magnifications which needs to be analyzed and semantically indexed in order to fully mine the potential knowledge held within the images.

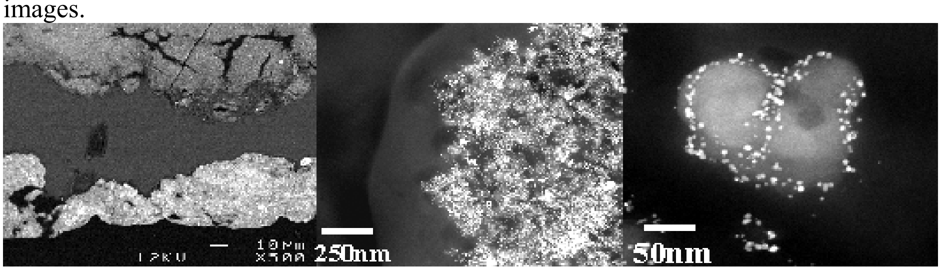

Fig. 1. Microscopic images of a fuel cell at 3 different magnifications

The remainder of this paper describes in more detail the semantic indexing system which we built and its application and evaluation within the context of the fuel cell domain. Although we have chosen one particular community in which to test our approach, the design is such that it applies to any image-analysis domain e.g., remote-sensing, satellite or medical images. The paper is structured as follows:

- Section 2 describes previous research on the generation of semantic image descriptions from low-level features;
- Section 3 describes the overall architecture of our system and its main components:
    - The ontologies which we developed;
    - The automatic image analysis programs which we invoked to extract low-level features and to generate MPEG-7 descriptions;
    - How we applied semantic inferencing rules to image data and the challenges this involved;
- Section 4 describes the Rules-By-Example graphical user interface which we developed;
- Section 5 contains empirical results obtained by testing our approach on real fuel cell images;
- Section 6 contains our conclusions and describes future plans for this research.

2 Previous Work

Most recent research in image or video indexing and retrieval has focused on query-by-example (QBE) [1-3]. However semantic querying or query-by-keyword (QBK) has recently motivated research into semantic indexing of images and video content. A number of research efforts have investigated the use of automatic recognition techniques to extract different low-level visual or audio features which together can be used to generate improved semantic descriptions of multimedia content. Recent attempts include work by Naphade et. al. [4] which proposes a statistical factor graph framework to bridge the gap between low-level features and semantic concepts. Chang et. al. [5] use a library of examples approach, which they call semantic visual templates. Adams et. al. [6] manually annotate atomic audio and video features in a set of training videos and from these develop explicit statistical models to automatically label the video with high-level semantic concepts. Zhao and Grosky [7] employ a latent semantic indexing technique which integrates a variety of automatically extracted visual features (global and sub-image colour histograms and anglograms for shape-based and colour-based representations) to enable semantic indexing and retrieval of images.

Overall the use of machine learning techniques to bridge the semantic gap between image features and high-level semantic annotations provides a relatively powerful method for discovering complex and hidden relationships or mappings. However the 'black-box' method often employed can be difficult to develop and maintain as its effectiveness relies upon the design and configuration of multiple variables and options. In addition, extensive and detailed training corpuses are required to ensure optimum performance from the system. The main disadvantage of machine learning-based systems is that they are built specifically for one particular domain and document type (e.g., images or speech) and cannot easily accommodate other domains or information sources.

In their recent paper, Marques and Barman [8] used machine-learning techniques to semantically annotate images with semantic descriptions defined within ontologies. Our approach is to replace or augment the machine-learning component with domain-specific inferencing rules defined by domain-experts through an intuitive user-friendly interface. Hatala and Richards [9] applied a similar approach to generating semantic descriptions for learning objects - relevant values for a particular metadata field are suggested by applying ontologies and rules – but they have not applied this approach to image content.

We believe that our approach enables some of the limitations in existing image annotation approaches (such as difficulty determining the distinguishing features and adapting to different domains) to be overcome. We do this through the complementary use of semantic web technologies and an interface which allows domain experts to intuitively and interactively develop and define semantic inferencing rules in an interoperable, machine-understandable and shareable format.

3 System Architecture

Figure 2 illustrates the overall system components and architecture for the knowledge management system that we developed to manage the manufacturing, performance, microscopy and image data captured from the fuel cell components.

The MPEG-7 (Multimedia Content Description) [10] standard is used to describe the low-level image features. The OME (Open Microscopy Environment) [11] standard is used to capture associated microscope details and settings – essential provenance information within any scientific context. An additional metadata schema (FUSION [12]) was developed specifically to satisfy the descriptive requirements of fuel cell analysts. We developed both XML Schemas [13] to define and validate the metadata descriptions, and OWL [14] ontologies, to define the semantics and semantic relationships between the terms used in each of these schemas. User-specified RuleML rules define the relationships between low-level automatically-extracted MPEG-7 features and high-level FUSION concepts or terms and are applied by an inferencing engine. All of the generated and validated metadata is stored in a central knowledge repository. A query engine, visualization engine and knowledge capture (annotation) tools sit on top of the knowledge repository. Together these enable users to access, interpret, assimilate and mine the stored data.

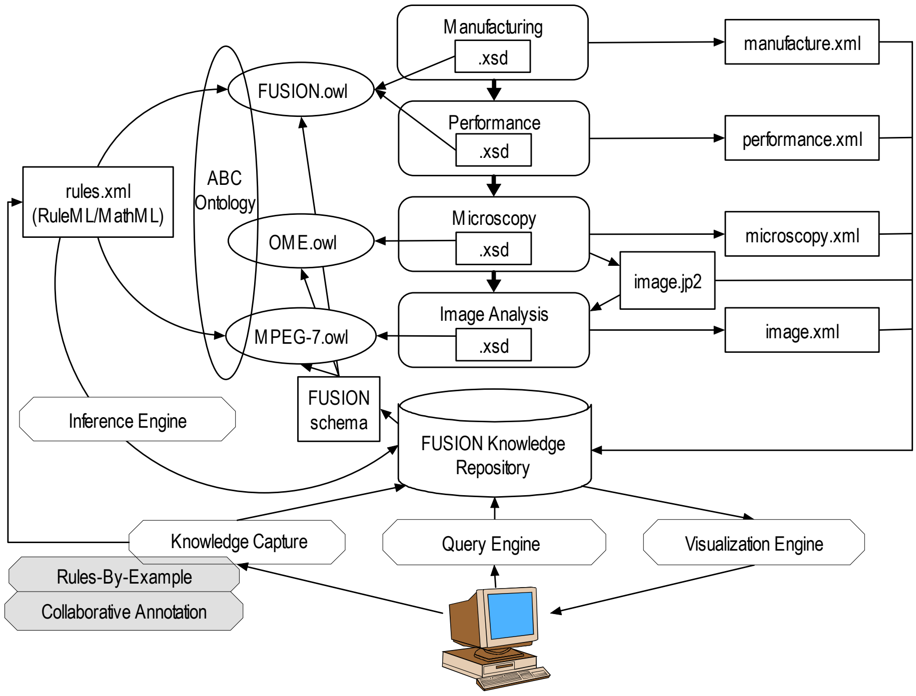

Fig. 2. Overall Architecture

A key aspect of our approach is the separation (as far as possible) of rules from facts. The rules are contained within explicit definitions in the RuleML files (rules.xml) and within implicit class/property relationships of the ontologies. Together these produce rules which are more flexible and able to be easily abstracted e.g., if a given rule applies to mpeg7:Color then it will also apply to mpeg7:DominantColor. The facts are built from the knowledge repository elements which are defined by both XML Schema (syntactic) definitions and ontological (semantic) definitions. Applying the rules through a reasoning engine, produces more "facts" which are inserted into the knowledge repository and defined within the related XML Schema and OWL ontology.

In order to enable semantic interoperability between the MPEG-7, OME and FUSION metadata vocabularies, and to define the semantic and inferencing relationships between the terms in these different schemas, we needed to develop ontologies for each of these vocabularies and harmonize or relate them. We used the top-level or core ABC ontology [15] developed within the Harmony project to do this. The ABC ontology provides a global and extensible model that expresses the basic concepts that are common across a variety of domains and provides the basis for specialization into domain-specific concepts and vocabularies. Brief descriptions of the MPEG-7, OME and FUSION ontologies together with figures illustrating a subset of these ontologies are provided below.

3.1 The MPEG-7 Ontology

Figure 3 illustrates graphically a subset of the MPEG-7 OWL ontology [16] which we developed to define the semantics of terms used within the MPEG-7 standard for describing multimedia/image content – and in particular low level visual features. Figure 3 shows only those classes for which "color" is a visual descriptor and the five sub-properties of color. A complete description of the MPEG-7 ontology is available at [17].

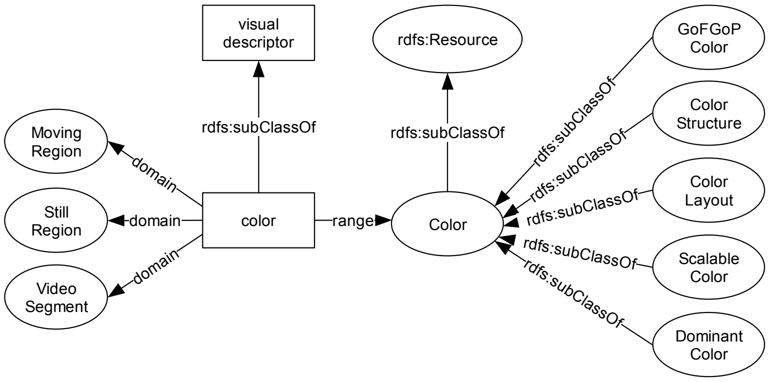

Fig. 3. The MPEG-7 Visual Descriptor "Color".

3.2 The OME Ontology

The OME (Open Microscopy Environment) [18] is an open source collaborative software project which aims to develop a common environment for the analysis, management and exchange of biological microscopic images. A key component of the OME is an XML-encoded file standard for the storage, analysis and exchange of image data output from microscopy procedures. A common standard for recording microscope-related metadata is essential for capturing the provenance or source of microscopy images and it includes such information as: the microscope manufacturer, serial and model, instrument settings, detectors, filters and light sources. An OME ontology was developed from the OME CoreSemanticTypes [19]. It defines the semantics associated with microscope output data and settings and enables them to be related to the image descriptors defined by MPEG-7 and FUSION ontologies. Figure 4 illustrates a subset of the OME OWL ontology that we developed.

3.3 The FUSION Ontology

Figure 5 illustrates the key classes or concepts and their relationships, as defined in the FUSION ontology that we developed for the fuel cell community. This was developed in collaboration with domain experts using the approach recommended in [20] and was used to generate the high-level semantic descriptions or terms which the fuel-cell experts use when searching and retrieving images.

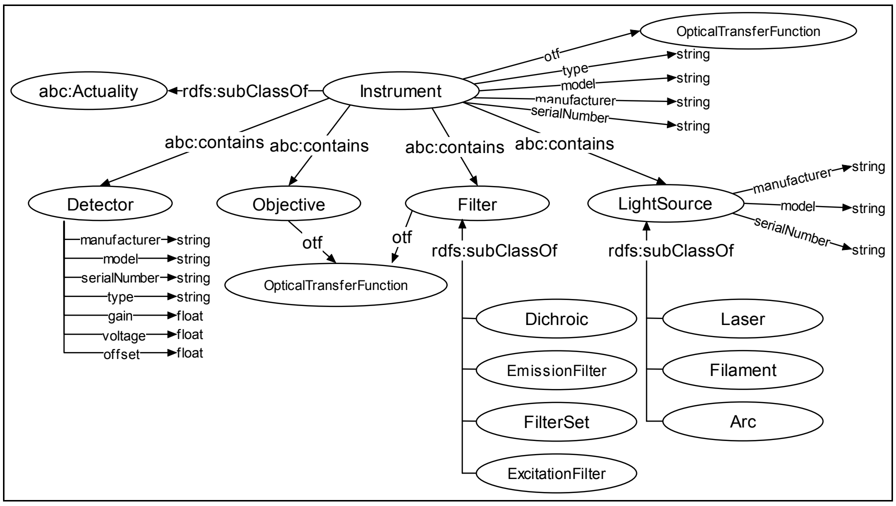

Fig. 4. A subset of the OME ontology showing an instrument and its components

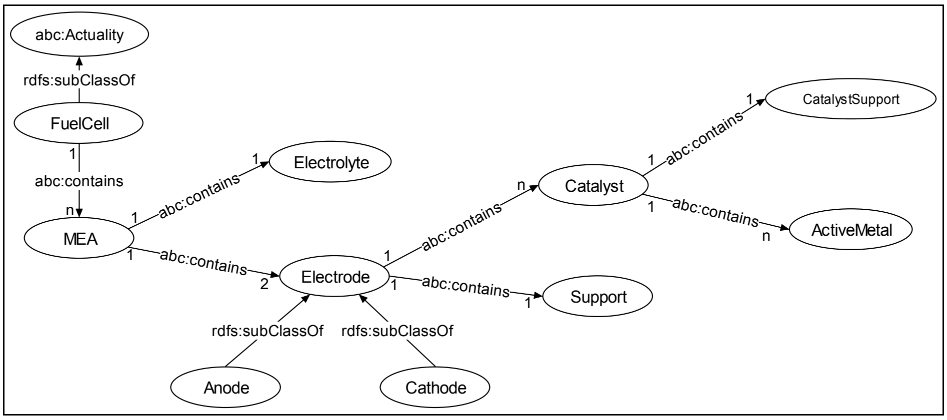

Fig. 5. A subset of the Fuel Cell ontology showing components of a fuel cell.

3.4 Automatic Feature Extraction

Automatic feature extraction for images [21] is a significant research area and a wide variety of tools and mechanisms exist for analyzing scientific and other types of images (e.g. [22]). Within the scope of this project, we are not particularly interested in the underlying method by which these programs process images but we are interested in the outputs or features that such programs produce and whether they can be used to generate MPEG-7 descriptions from which higher-level semantic concepts can be inferred. For our system we have chosen to use MATLAB [23] because it is a popular and powerful tool which is currently being widely used by microscopists and scientists for analyzing scientific images. MATLAB is capable of producing a large amount of low-level data about the features and objects within an image e.g., area, location, pixel colour, centroid, eccentricity, orientation etc. Currently we are calling MATLAB analysis methods directly but plan to employ SOAP [24], WSDL [25] and

OWL-S [26] to make the automatic feature extraction modules of MATLAB available as web services. This would allow greater flexibility and adaptability by enabling the optimum extraction and analysis tools or services to be plugged in at runtime. The data produced by MATLAB analysis methods is transformed to MPEG-7 descriptions using Python scripts.

3.5 Defining and Invoking Semantic Inferencing Rules

Many challenges need to be overcome, when implementing inferencing rules within a Semantic Web application. [27] provides a useful discussion of the current approaches to rules within the Semantic Web and examines some of the motivations and difficulties. The three main issues faced by our project were: what recording syntax to choose; how to build or generate the rules; and how to apply the rules.

The recording syntax had to be flexible, machine-processable and interoperable with existing tools and standards (e.g., XML Schema, OWL). By using an XML-based rule definition language such as RuleML, the rules, ontologies and metadata schemas can be linked through the use of URIs, namespaces and semantic attributes defined in the XML Schema documents (see [28] for further details on this approach). In addition, as noted by Boley et. al. [29], XML-based rules provide a sharable, application-independent form of knowledge which can easily be processed for each application using transformation methods such as XSLT or others of the growing variety of generic tools for the creation, modification and management of XML data. Several possible XML formats exist for representing inference rules. The most popular markup language for rules and the current most likely candidate for adoption by the Semantic Web community is RuleML [30], the Rule Markup Language. Although still under active development, RuleML has a wide user base, a reasonable level of available tool support in the form of reasoning engines and transformation options and provides an XML description of inference rules. Alternative markup languages which were examined included the eXtensible Rule Markup Language (XRML) [31] and the Description Logic Markup Language (DLML) [32] both of which lacked the breadth of users and tool support available for RuleML. A promising emerging approach, the Semantic Web Rule Language (SWRL) [33, 34] which combines OWL and RuleML, is also under-development and may be considered for use within this project as its stability and support base expands. In addition MathML [35], the W3C's Mathematical Markup Language, was also investigated in order to support the description of more complex mathematical relationships and models relating to the fuel cell domain.

While XML-based rules have many advantages, they can be very difficult and time consuming to create and in the case of image data, rely on the specification of low-level features which can be highly complex (e.g., colour histograms). The size of the resulting XML file is also a serious issue. For example, the simple rule provided in the introduction to this paper would convert to approximately 40 lines of RuleML. The translation from rules specified in a natural language format to RuleML is not necessarily a one-to-one mapping. For example, *color is 'one of' set* can have multiple possible RuleML representations. In order to assist domain experts in the creation, refinement and application of the semantic inferencing rules, we developed the Rules-By-Example GUI – an intuitive user interface generated from the configured backend ontologies that allows users to build complex RuleML rules quickly and easily. This is described in detail in the next section.

Once the rules have been defined and recorded they need to be processed and applied to the specified data using an inferencing engine. Possible candidates for the inferencing engine include: JESS (Java Expert System Shell) [36]; and Mandarax [37], a Java RuleML engine. Problems that often arise when invoking the semantic inferencing rules include: slow processing speed; the need to convert data to in-memory RuleML facts and the lack of native RuleML support for applying standard string and mathematical relations (e.g., greater than, equal to, etc.). Mandarax provides some internal representation and processing options (ZKB) which can be used to overcome these issues. Alternatively using wrapper scripts and relational or XML databases can also be used to improve performance and flexibility.

4 Rules-By-Example Implementation and Interface

Figure 6 illustrates the dataflow within the Rules-By-Example (RBE) system. The RBE GUI enables the user to specify values for visual features from palettes of sample colors and textures, drawing tools or by specifying example images selected from the knowledge repository. For example, using the graphical tools, users can easily specify that 'color' should be 'like this region in this image'. By generating the GUI's pull-down menus from MPEG-7 descriptions generated by the image analysis services and the ontologies specified at configuration time, the GUI enables domain experts to easily define complex rules and store them in RuleML format without understanding the complexities of image analysis or RuleML.

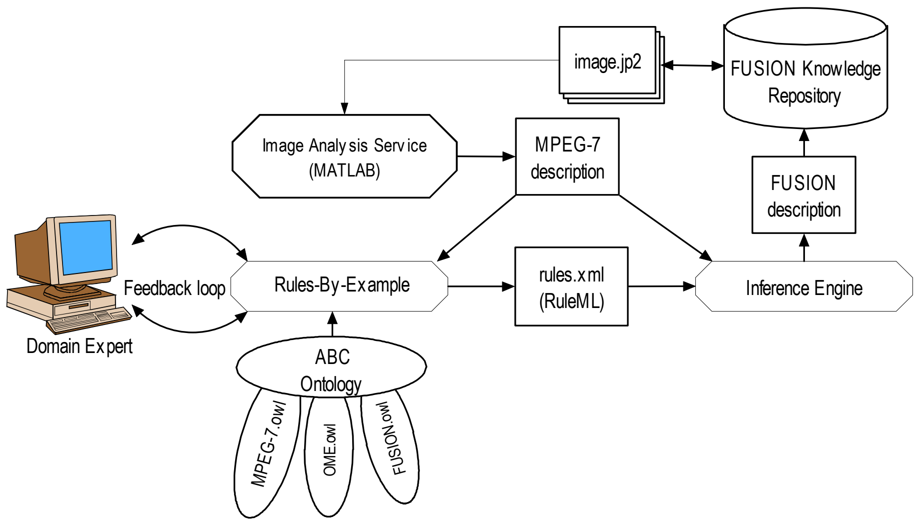

Fig. 6. Rules-By-Example Dataflow

Figure 7 illustrates a screenshot of the prototype RBE interface developed. The menu options within the interface are populated dynamically from the back-end ontologies specified during system configuration. For example, the initial menu (marked A) identifies the MPEG-7 Media element to which the rule applies. Consequently the drop-down menu displays subclasses of the Region class of the MPEG-7 ontology. In this case the user has selected the StillRegion sub-class. Further use of relationships derived from ontologies can be seen in the menu marked B. Depending on the media element selection made in menu A, menu B displays all possible properties (or attributes) which have that element or its superclass as the domain. The possible operators and values for a property, are also generated from its range definition e.g., if the type is *string* then relationships such as *less than* are irrelevant. This type of dynamic population of the user interface provides both user support and domain independence – both strengths of this system. Standard logical operators (AND/OR) enable the specification of more advanced, composite, logic-based rules. The window in the upper right-hand section of Figure 7 shows the metadata which has been automatically extracted for the highlighted region.

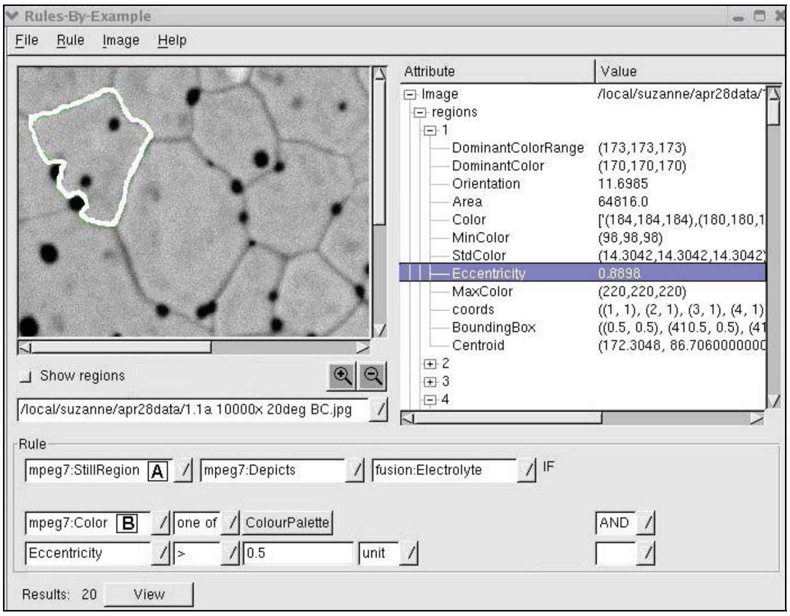

Fig. 7. Screenshot of user interface prototype

In order to specify the values of atoms within a rule's body, the user is able to select an existing image (or set of images) and choose regions or segments of the image to use as examples. For instance, in the screenshot above, the color specification within this rule states that for the rule to be true (that is for the *StillRegion* to depict *Electrolyte*) then color must be one or more of the set shown on the palette. The palette displays all of the colors in the selected segment and the user is able select one or more. An additional complexity when searching for visual features is the fuzzy matching problem. Similarity rather than exact matches are usually required. To support this, the interface provides the ability to set a threshold value for individual image attributes or features.

Once the user is satisfied with a rule definition, it can be transformed and saved in RuleML format (possibly augmented with MathML). The RuleML file generated from the rule in Figure 7 is available[1]. Manual creation of such a rule would have been extremely difficult and time consuming.

Overall the Rules-By-Example interface allows the user to quickly develop, apply and refine highly complex rules without needing to understand complex low-level MPEG-7 terms or values. In addition, it enables users to direct the system so that it focuses on the objects, regions or distinguishing features of highest priority or interest – in contrast to traditional approaches which require pre-selection of important features. The next section discusses the results of initial system evaluation.

5 System Evaluation

5.1 Evaluation Process

Initial evaluation tests were carried out using a set of sample fuel cell images. The objectives of these experiments were: to determine how the semantic labels generated by our rules-based system compare with manual annotations; and to gather feedback on the usability and expediency of the RBE interface.

The evaluation process was conducted as follows:

1. Twelve microscopy images, of equivalent magnification (10000x) and from the same fuel cell component (electrolyte of a ceramic fuel cell) were manually segmented by defining region boundaries. The images were all grayscale and either high quality JPEG or TIFF images of average size 2560x1920 pixels. Two main region types were present: large pale gray 10% yttria-stabilized zirconia (10YSZ) grains and smaller, darker alumina grains.
2. The images were analyzed using the MATLAB image processing toolbox and low level data (including: Boundary Coordinates, Area, BoundingBox, Eccentricity, Color (lightest, darkest, mean, standard deviation, dominant color, dominant range of colors)) were extracted and saved to a MySQL database.
3. The RBE interface was used by the system designers to manually assign labels to all of the regions in the images (total of 274 regions).
4. The RBE interface was then used by the fuel cell experts to build six sample RuleML rules (three describing the 10YSZ grains and three describing the alumina grains) based on a region's Area, Eccentricity, DominantColor and DominantColorRange. Feedback on the usability of the interface was recorded during this phase.
5. A Python script converted the data in the database into RuleML facts describing the features of each region. Some filtering and preprocessing of the facts was necessary to remove tiny (<100 pixel) areas that were detected by the segmenting software but which do not represent relevant regions. These regions can greatly reduce the speed of the inferencing engine. Secondly, as discussed in section 3.5, mathematical relationships defined within the rules (e.g., *greater than*) are not

[1] http://metadata.net/sunago/fusion/sample.ruleml

processed by Mandarax. So in order for the inferencing engine to apply rules such as, “DominantColor must be greater than (165,165,165)”, a Python pre-processing script was required to generate additional facts e.g., declaring those regions with attribute values greater than the nominated threshold.

6. A sample RuleML query was defined and the rule, query and facts were combined into a rulebase and processed using the Mandarax inferencing engine. The list of regions which matched the query were saved and compared with the labels of the manually annotated regions to calculate the precision and recall of the RBE system.

5.2 Evaluation Results

Table 1 shows the results of initial evaluation tests generated by applying the six rules specified below the table. Initial results demonstrate exceptional precision but only mediocre recall. The high precision is a direct result of the RBE interface which enables users to interactively adapt and refine their rules and filter out erroneous rules, based on system feedback. However this also encourages users to mould rules to fit the test set, potentially at the expense of performance or accuracy when applied to a larger size or range of images. The improvement in recall shown when two rules were combined (Rules 3 and 6) demonstrates how this tendency can be overcome and leveraged to improve the results.

Table 1. Summary of initial evaluation results

	1	2	3	4	5	6
Total regions which match the rule (I)	66	52	102	24	24	33
Regions which match the rule and have the equivalent label in the database (H)	66	52	102	23	24	32
Total regions labeled as *component* in the database (C)	167	167	167	53	53	53
Precision (C/I)	100	100	100	95.83	100	95.83
Recall (C/H)	39.52	31.14	61.08	43.4	45.28	60.38

1. mpeg7:Region mpeg7:Depicts 10YSZ IF mpeg7:DominantColor is ‘(1,1,1)’ AND DominantColorRange is ‘(6,6,6)’
2. mpeg7:Region mpeg7:Depicts 10YSZ IF Area < 5000 and Eccentricity < 0.5
3. mpeg7:Region mpeg7:Depicts 10YSZ IF Rule1 OR Rule2
4. mpeg7:Region mpeg7:Depicts alumina IF mpeg7:DominantColor is > ‘(165,165,165)’ and mpeg7:DominantColor < ‘(200,200,200)’
5. mpeg7:Region mpeg7:Depicts alumina IF Area > 100000 and Eccentricity > 0.5
6. mpeg7:Region mpeg7:Depicts alumina IF Rule4 OR Rule5

The homogeneous nature of the images also contributed to the high precision of the results. The images were highly consistent because they originated from the same fuel cell component (albeit at different cross-sections) and were created under exactly the same microscopy conditions. This consistency, together with the presence of only two distinct region types, made the creation of precise rules relatively simple.

However, the slow performance of the system was a serious problem. The size of the images, the number of regions and the size of the associated data (100MB text files generated by MATLAB) significantly affected the performance and responsiveness of the RBE interface. Processing even a simple query and rule over a large set of facts took up to 10 minutes. A number of solutions (e.g., closer integration of database functionality, use of MATLAB APIs and faster image processing techniques) are currently being investigated.

User's feedback on the interface produced several useful requests including: the ability to compare and contrast selected sample regions; a 'sketch' interface integrated with feature measurements such as *eccentricity*, to assist in describing shape; and enabling the definitions of the ontology terms used to populate the dropdown menus to be displayed. These interface issues are currently being addressed.

6 Future Work and Conclusions

The initial test results were very encouraging. A comparison of the semantic labels generated automatically (by applying user-defined rules) with the semantic labels manually attached to the same images, has demonstrated very high correlation. Given the semantic descriptions generated by applying the rules defined within the RBE GUI, domain-experts are able to perform much more complex and sophisticated queries over the image data, in terms which are useful and familiar to their domain e.g., "give me all of the images depicting Catalysts with high ActiveMetal density".

However further testing is required – using larger image sets and images from other domains. Future tests will also involve more formal interface evaluation (user interviews, observation, task analysis) and more comprehensive investigation of methods for improving performance and scalability. We also plan to compare the performance of the system with machine learning based methods for image analysis.

Additional future plans include: supporting user suggestions gathered from the initial trials; including confidence ratings with rule specifications; extending the approach described here to the semantic indexing of video content (exploiting automatically-extracted audio and moving object features); adding an annotation component to allow shared evaluation of experimental data.

To conclude, we have described in this paper a new approach to the semi-automatic semantic indexing of images. Rules-By-Example (RBE) combines RuleML with a Query-By-Example type interface to enable domain-experts to graphically define domain-specific rules that can infer high-level semantic descriptions of images from combinations of low-level visual features (e.g., color, texture, shape, size of regions). By applying these rules, the system is able to analyze the visual features of any given image from this domain and generate semantically meaningful descriptions, using terms defined in the domain-specific ontology. We believe that the advantages of this approach include:

- Faster, more accurate cost-effective semantic indexing of images;
- A graphical user interface which provides an intuitive, user-friendly means by which domain-experts can interactively develop, apply and refine semantic inferencing rules for images;

- A flexible architecture which enables any ontology to be easily linked to the system at configuration time, dynamically customizing the system for a particular domain;
- The ability to plug in the optimum or most appropriate automatic image analysis tools/services at runtime;
- The use of machine-understandable XML-based image metadata standards, schemas, ontologies and rule languages which ensure maximum interoperability of the system and data;
- A knowledge capture system which accurately and easily captures the knowledge of domain experts and maximizes the potential of their image content to be discovered, re-used and assimilated by communities, services, tools and agents on the Web.

Acknowledgements

The work reported in this paper has been funded in part by the Co-operative Research Centre for Enterprise Distributed Systems Technology (DSTC) through the Australian Federal Government's CRC Programme (Department of Education, Science and Training). Thanks also to staff from the Centre for Microscopy and Microanalysis at the University of Queensland for their feedback and the images used in this project.

References

1. Deng, Y. and Manjunath, B. S. "Content-based search of video using color, texture and motion." in International Conference on Image Processing (ICIP '97). Washington, DC, USA, October 26-29, 1997.
2. Marchand-Maillet, S., "Content-Based Video Retrieval: An Overview." University of Geneva, Geneva, Switzerland. 2000. Technical Report No. 00.06, CUI.
3. Zhang, H., Wang, A., and Altunbasak, Y. "Content-based video retrieval and compression: a unified solution." in IEEE International Conference on Image Processing. Santa Barbara, CA, Oct 1997.
4. Naphade, M. R., et al. "A Factor Graph Framework for Semantic Indexing and Retrieval in Video", in IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'00). Hilton Head, South Carolina, June 16 2000.
5. Chang, S. F., Chen, W., and Sundaram, H. "Semantic Visual Templates: linking visual features to semantics", in IEEE International Conference on Image Processing (ICIP '98). Chicago, Illinois, 1998.
6. Adams, B., et al., Semantic Indexing of Multimedia Content Using Visual, Audio and Text Cues. EURASIP Journal on Applied Signal Processing. 2003.
7. Zhao, R. and Grosky, W. I., Negotiating The Semantic Gap: From Feature Maps to Semantic Landscapes. Pattern Recognition. 35(3): p. 51-58, 2002.
8. Marques, O. and Barman, N. "Semi-automatic Semantic Annotation of Images Using Machine Learning Techniques", in International Semantic Web Conference. Florida, October 2003.
9. Hatala, M. and Richards, G. "Value-added Metatagging: Ontology and Rule based Methods for Smarter Metadata", in Rules and Rule Markup Languages for the Semantic Web (RuleML2003), 2003.
10. ISO/IEC 15938-5 FDIS Information Technology, "Multimedia Content Description Interface - Part 5: Multimedia Description Schemes". MPEG7, Sydney. July 2001.

11. Open Microscopy Environment, OME, (http://www.openmicroscopy.org/).
12. FUSION Project, (http://metadata.net/sunago/fusion.htm).
13. W3C, XML Schema Specifications (Part 0: Primer, Part 1: Structures, and Part 2: Datatypes), W3C Recommendation, May 2001, (http://www.w3.org/XML/Schema#dev).
14. W3C, OWL Web Ontology Language Overview, W3C Recommendation, 10 Feb 2004, Edited by Deborah L. McGuinness and Frank van Harmelen, (http://www.w3.org/TR/owl-features/).
15. Lagoze, C. and Hunter, J., The ABC Ontology and Model (v3.0). Journal of Digital Information. 2(2), 2001.
16. Hunter, J. "Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology", in International Semantic Web Working Symposium (SWWS). Stanford, July 30 - August 1, 2001.
17. MPEG-7 Ontology, (http://metadata.net/harmony/MPEG7/mpeg7.owl).
18. Swedlow, J. R., et al., Informatics and Quantitative Analysis in Biological Imaging. SCIENCE. 300: p. 100-102, 2003.
19. Open Microscopy Environment, OMECoreSemanticTypes.ome, 4 Sep 2003, (http://openmicroscopy.org/XMLschemas/STD/RC2/OMECoreTypes.ome).
20. Noy, N. and McGuinness, D., "Ontology development 101: A Guide to Creating Your First Ontology". Stanford University, 2001. Technical Report SMI-2001-0880.
21. Nixon, M. and Aguado, A., Feature Extraction and Image Processing. 2002: Newnes.
22. Internet Analysis Tools Registry, (http://www.cma.mgh.harvard.edu/iatr/display.php?search_type=all).
23. MathWorks, MATLAB, (http://www.mathworks.com/products/imageprocessing/).
24. W3C, SOAP Version 1.2 Part 0: Primer, W3C Recommendation, 24 June 2003, Edited by Nilo Mitra, (http://www.w3.org/TR/2003/REC-soap12-part0-20030624/).
25. Christensen, E., et al., Web Services Description Language (WSDL) 1.1, W3C Note, 15 March 2001, (http://www.w3.org/TR/wsdl).
26. The DAML Services Coalition, OWL-S: Semantic Markup for Web Services, May 2003, (http://www.daml.org/services/owl-s/1.0/).
27. Staab, S., et al., Trends & controversies - Where are the rules? Intelligent Systems, IEEE. 18(5): p. 76, 2003.
28. Hunter, J. and Lagoze, C. "Combining RDF and XML Schemas to Enhance Interoperability Between Metadata Application Profiles", in WWW10. HongKong, May 2001.
29. Boley, H., Tabet, S., and Wagner, G. "Design Rationale of RuleML: A Markup Language for Semantic Web Rules", in Semantic Web Working Symposium (SWWS), 2001.
30. The Rule Markup Initiative, RuleML, (http://www.dfki.uni-kl.de/ruleml).
31. eXtensible Rule Markup Language, XRML, (http://xrml.kaist.ac.kr/).
32. Description Logic Markup Language, DLML, (http://co4.inrialpes.fr/xml/dlml/).
33. Horrocks, I., et al., "SWRL: A Semantic Web Rule Language Combining OWL and RuleML". 19 November 2003. proposal
34. Grosof, B. N., et al. "Description Logic Programs: Combining Logic Programs with Description Logic", in 12th Intl. Conf. on the World Wide Web (WWW-2003). Budapest, Hungary, May 20-23, 2003.
35. W3C, Mathematical Markup Language (MathML), W3C Proposed Edited Recommendation, 04 August 2003, Patrick Ion Edited by David Carlisle, Robert Miner and Nico Poppelier, (http://www.w3.org/TR/2003/PER-MathML2-2003080).
36. Friedman-Hill, E., Jess, The Expert Shell for the Java Platform, (http://herzberg.ca.sandia.gov/jess/).
37. Dietrich, J., The Mandarax Project, (http://www.mandarax.org/).