Media Analysis and the Algorithm Ontology

Patrizia Asirelli[1], Sara Colantonio[1], Suzanne Little[2]
Massimo Martinelli[1] and Ovidio Salvetti[1]

[1] Institute of Information Science and Technologies (ISTI)
Italian National Research Council (CNR)
Via Moruzzi 1, 56124, Pisa, Italy

[2] Institute of Computer Vision and Applied Computer Sciences (IBaI)
Leipzig, Germany

Abstract. Media analysis algorithms are used for a variety of purposes. They may improve media facets such as contrast or signal-to-noise ratio or extract low-level details such as MPEG-7 features to be used in data mining and other higher-level processing. However, algorithms are difficult to manage, understand and apply in particular for non-expert users. Therefore we are developing an algorithm ontology to support identification, aggregation and recording of algorithms for media analysis. This is especially useful for domains with high-volumes of complex media objects to investigate and integrate. Algorithms for media analysis may be applied at multiple points within a typical multimedia lifecycle. This article discusses a proposed algorithm ontology to support identification, retrieval and application of multimedia analysis processes and its application to metadata management and multimedia interoperability.

1 Introduction

Advances in tools and technologies for digital media production and analysis have brought the need for more complete and interoperable descriptions of media processing to the forefront. Information captured at each stage of the multimedia lifecycle is of great value for tasks such as analysis, data mining and media reuse.

For example, in the medical research field media data is a common output of the experimental evaluation phase where specimen or patient data consists of numerical measurements and scans, micrographs or visualizations. This media is often normalized, segmented and analyzed using a variety of media analysis algorithms and then integrated into the larger pool of data for investigation and evaluation. This is also true of other scientific research fields and areas such as industrial monitoring or digital art curation where media is captured and processed before being used to assess a process, mined for knowledge or to evaluate a theory.

Therefore, the media that is produced has a variety of possible applications and often undergoes post-production processing and possibly further low-level analysis procedures before being annotated, evaluated or applied. Even once media has been as-

Asirelli P., Colantonio S., Little S., Martinelli M. and Salvetti O. (2008).
Media Analysis and the Algorithm Ontology.
In *Image Mining Theory and Applications*, pages 16-25
DOI: 10.5220/0002341300160025

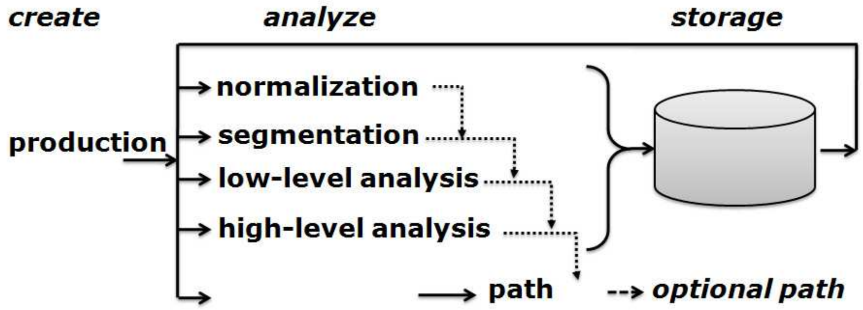

Fig. 1. Three main phases of media processing and analysis: Create, Analyze, Storage. Multiple, potentially iterative paths are possible.

similated into the general data set it may still undergo analysis processes at later stages within its life cycle.

Figure 1 shows an abstract illustration of a possible media process. Firstly, algorithms may be applied in pre-processing directly after the capture of the media. Examples of algorithms used here may include those to increase contrast, improve signal-to-noise ratio or normalize variations.

Secondly, algorithms may be used in media analysis possibly for the purpose of data mining or fine-grained semantic annotation. This generates a set of low-level analysis data from the media object. Algorithms used here may include segmentation or edge detection algorithms followed by processes to extract low-level features such as those defined by the MPEG-7 standard. In addition, higher level processing using machine learning techniques such as neural networks [19] or case-based reasoning [25] and approaches such as semantic inferencing rules [17] may be used to derive semantic annotations from the low-level details.

Thirdly, algorithms may be used for media conversion or touching up prior to reuse of the media or applying secondary analysis processes. This will usually include algorithms similar to pre-processing but potentially occurs at multiple later stages throughout the media lifecycle. Examples may include reducing the size or resolution of an image, format conversion or specific processing prior to applying 3D volume rendering or isosurface generation.

There are a number of reasons why a clear, formal description of processes, algorithms or methods applied to media objects is a useful and necessary part of multimedia metadata and the multimedia lifecycle. These descriptions should be detailed – not only general process descriptions but specific definitions of the requirements, formats and outcomes relating to media analysis algorithms.

Firstly, clear definitions of algorithms are useful for the identification of syntactically appropriate algorithms. For example, algorithms that require input media to be of colour type RGB as opposed to binary or black and white images.

Secondly, higher-level semantic descriptions enable the use of pre-existing examples or case-studies to develop solutions for similar problems. For example, using broad general statements of the final goal such as "segment this image" or "improve image *quality*” and developing a sequence of possible processes to apply to achieve the desired outcome For example, a researcher needs to reduce the noise and improve the contrast in a radiology image prior to analysis and interpretation but is unfamiliar with the specific algorithms that could apply in this instance.

Finally, it is important to keep an accurate, complete and defensible record of the processing that has been applied to a media object within its entire lifecycle. This type of provenance data is especially important in scientific or investigative domains. In addition, many applications require the processes applied to media to be concisely recorded for re-use, re-evaluation or integration with other analysis data.

The problem is that algorithms for media analysis are difficult to manage, understand and apply, particularly for non-expert users. The main difficulty lies in quantifying and articulating the “visual” (or “aural”) result of an algorithm so that its purpose and outcome can be unequivocally understood and interpreted independently of the media domain.

Therefore, we are developing an algorithm ontology that aims to record and describe available algorithms for application to image (and eventually other media) analysis. This ontology can then be used to interactively build sequences of algorithms to achieve particular outcomes. In addition, the record of processes applied to the source image can be used to define the history and provenance of data.

This article presents an outline of the algorithm ontology, its use within an example scenario and discusses how it can be applied to multimedia metadata management and to promote interoperability of multimedia metadata.

2 Related Work

The multi-dimensional nature of multimedia metadata and the challenges this presents when integrating media, particularly in a web-based system, is a well-known problem [9], [26]. A large number of initiatives aiming at standardizing metadata have come to light in recent years to describe multimedia content in different domains and to enable sharing, exchanging and interoperability across a wide range of networks. According to their functionality, two types of standards can be distinguished:

- One is directly related to the representation of multimedia content for a specific domain and provides a standardized description scheme or well-defined syntax.
- The other integrates metadata standards from different domains to provide metadata models or broad semantic definitions and enhance general semantic interoperability.

A variety of standards to describe and define multimedia objects and their contents have been proposed such as MARC [21], Dublin Core [13]), VRA Core [28], LOM [20], DIG35 [12], MPEG-7 [22] and MPEG-21 [23]. A general comparison and review of these standards can be found at [24]. These standards are illustrative of the first type. However, they generally lack sufficient detail to describe low-level media features and tend to concentrate on more abstract metadata and semantics.

Within these standards, the use of MPEG-7 for the description of multimedia documents is wide-spread. MPEG-7, named “Multimedia Content Description Interface”, is a standard for describing the multimedia content data. It is not aimed at any one application in particular; rather, the elements that MPEG-7 standardizes support as broad a range of applications as possible. MPEG-7 provides a document description language (DDL) to encode a structured, schema-based model to describe media-specific properties of audio, video, and text data, as well as the individual content objects within each primitive media stream.

Within the second type of standard a number of general multimedia ontologies have been proposed to support the interoperability of multimedia tools and metadata. A number of these are based on the MPEG-7 standard and define the key concepts as classes and properties. For example, Hunter [16], Tsinaraki et al. [27], and AceMedia [1] are all based on the MPEG-7 standard.

Semantic definitions of algorithms can also be found in the algorithm pattern of the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [14] and in an ontology proposed in [2] where structures for detecting, classifying or annotating a region of an image are included with more generic media concepts. M-OntoMat-Annotizer [8] also includes concepts recording processing knowledge. A semantic framework proposed in [11] includes an analysis ontology that aims to describe abstract media processes. Finally a general thesaurus for image analysis purpose has been developed [4], [7]. All of these standards provide some level of semantic structure for defining media and analysis processes.

The difficulty is that these standards do not provide sufficient levels of detail to address multimedia understanding problems. For instance, extensions are required to the MPEG-7 standard to define specific low-level analysis features such as ‘eccentricity’, ‘color range’, etc. Previous work by Hollink et al. [15] describes some extensions to Hunter’s MPEG-7 ontology by creating subproperties of the visual descriptor to incorporate analysis terms.

In the end, we have both, to acknowledge the need to extend the available technology towards multimedia ontologies, and to add more semantics in order to be able to handle applications which require annotation, retrieval and summarization of multimedia documents.

3 Algorithm Ontology

Previous sections have shown that formal semantics describing media analysis algorithms are needed to address issues such as discovery of algorithms, choreography of algorithms and recording of provenance. This section describes the ongoing development of an ontology to describe and define image analysis algorithms and presents an example scenario illustrating how this ontology may be applied.

Starting from image understanding problems (e.g. segmentation, analysis, etc.), we are working to the realization of a thesaurus containing related concepts and algorithms. It would be important to represent features on how metadata are extracted so that they can be used at a higher level for images, and multimedia information handling in general. We have defined this meta-information as morpho-densitometric (shapes, how the object is made, etc.) and spectral characteristics.

Another important feature would be to record how a specific result has been obtained starting from a particular input to represent the set of algorithms and/or procedures used to produce a particular result. This work is also based on the preliminary results obtained in a collaborative project with the Dorodnicyn Computing Centre of the Russian Academy of Science that developed a technical vocabulary of more than 1000 terms describing image characteristics, algorithms used to obtain images, and relations among terms [3].

Figure 2 shows a class diagram of part of the algorithm ontology. The main concept is the Algorithm class which has a number of subclasses that classify the different types of algorithms (FilterAlgorithm, SegmentationAlgorithm, etc.). Information about each algorithm, such as the Input, Output, any Preconditions and the Effect are also included. The effect of applying an algorithm is the most difficult concept to articulate. This is a key area of research as we endeavor to define the outcome in an independent manner. The ontology can also be integrated with existing media ontologies, such as those referenced in section 2, to define the class and characteristics of media.

3.1 Example Scenario

To be more precise on the kind of problems we want to face we introduce the following scenario.

The Problem:
Classify the dense breast tissue in mammography images, according to the BI-RADS classification [18].

Hypothesis of solution:

Step 1: Get a digital mammogram of patient P (result: image I_0, see Figure 3).
Step 2: Improve the quality of the input image I_0 by applying a digital filter (result: image I_1).
Step 3: Extract the different tissues by applying a region-based segmentation algorithm to the input image I_1. As a result, image I_1 is partitioned in a set of N homogeneous regions (result: image $I_2 = \bigcup_{i=1,\ldots,N} \{R_i\}$).
Step 4: Select the region R_d corresponding to the dense tissue and describe it by applying a list of algorithms for computing geometrical and densitometric properties (result: array A).
Step 5: Classify the tissue by applying a classification function to the input array of features A (result: the label of the density class $C \in \{$Class I, Class II, Class III, Class IV$\}$).

However, for each of the previous steps a question arises regarding the choice of the different algorithms that can be applied to the input data.

In Step 2, the issue regards the digital filter selection: a number of filters, e.g. Fourier, Wiener, Smoothing, Anisotropic, etc., can be applied having different input-output formats and giving slightly different results. However, a filter that preserves edges, reduces noise and removes small curvilinear structures can be the best choice in the case at hand.

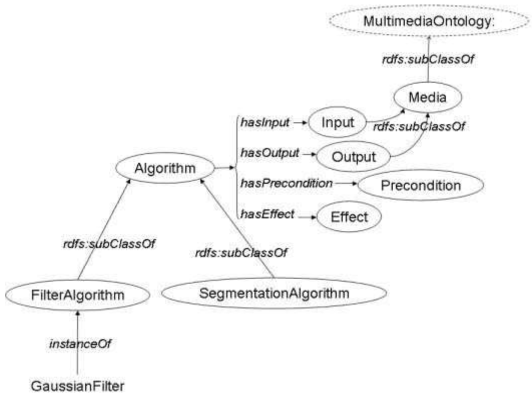

Fig. 2. Class diagram of part of the Algorithm Ontology.

In Step 3, several segmentation algorithms can be considered for region extraction: clustering, histogram, homogeneity criterion, etc.. The choice can heavily rely on the image type and the problem at hand.

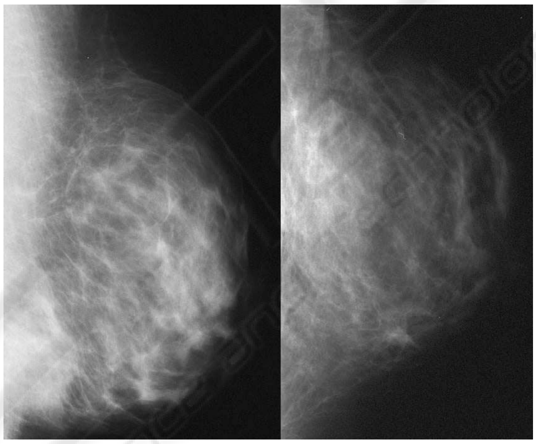

Fig. 3. Digital mammograms.

In Step 4, the question is related to the selection of significant geometrical and densitometric properties used for describing the extracted region. Usually, several possibilities are available, depending on the considered mathematical models for describing closed curves (regions) and the grey level distribution inside each region (histogram, Gaussian-like, etc.), but supplying the user a tool for selecting the most relevant one according to their meaning could reveal as one of the most viable solution.

In Step 5, the question regards the classification function, which can be defined according to one of the several pattern recognition methods. In many cases, such methods are adaptive and this further complicates the selection.

This type of detailed analysis processes where decisions made in each step have a significant impact upon choices in following steps and where multiple media objects or textual/numerical metadata may result are common in these types of scenarios.

4 Applying the Algorithm Ontology

4.1 An Infrastructure for Multimedia Metadata Management

MultiMedia Metadata Management (4M) Infrastructure [5], [6] has been developed as part of MUSCLE (Multimedia Understanding through Semantics, Understanding and Learning) through the EU Network of Excellence (NoE) initiative with the aim of supporting multimedia analysis, exchange and foster collaboration among research groups. The infrastructure consists of five main co-operating units, devoted to feature extraction from multimedia objects, database management, algorithms and annotations handling, and integration. 4M has been designed taking into account the use of the most promising existing tools, open-source software, java-based implementations and multimedia metadata standards.

The system has four main goals:

- To store, organize and retrieve distributed multimedia resources;
- To manage algorithms for information processing;
- To add semantic annotations;
- To access, protect and/or share information.

One possible processing sequence using the 4M architecture is as follows.

1. media upload;
2. media analysis to produce MPEG-7 feature data;
3. media storage with metadata in XML database;
4. identification of algorithms to achieve a user goal;
5. application of algorithms to media;
6. recording of outcomes in the database;
7. query of data to find related media;

The 4M infrastructure provides an environment for handling both low-level features extracted from multimedia objects, and semantic high-level information coming from automatic and semi-automatic processes of annotation and finally for managing, integrating and processing all of this information.

The algorithm ontology, within the 4M infrastructure, can assist in classifying acquired knowledge about a domain and help users to solve related problems. For example, in stages 2, 4 and 5 of the suggested process the algorithm ontology provides a standardised set of terms for searching, comparing and applying media processing functions.

Within the 4M infrastructure, the algorithm ontology enables users to query for available processing tools based on their classification (alg:Algorithm), on their effect (alg:hasEffect) or to browse for available processing options based on the current input data format (alg:hasInput). It also enables the recording and storage of media processing steps in a clear and independent manner.

4.2 Multimedia Interoperability

The algorithm ontology can also play an important role in the interoperability of multimedia metadata. This interoperability can be approached from two points of view: low-level (syntactic) interoperability and high-level (semantic) interoperability.

At the low-level, it concerns formats and data structures, eventual transformations between them and therefore the related algorithms of conversion. Sometimes an adequate and optimal conversion cannot ignore information related to data itself. Thus an ontology can be associated with the conversion process that can interest the algorithms (different but computationally equivalent algorithms); media associated to data structures and the data structures themselves (different but associated with the same media).

At the high-level, interoperability concerns the domain problem, that is the analogy of only apparently different or distinct problems but similar in reality and that therefore can be faced and solved with the same methodology (computational procedure). In this case, domain problem semantics can be codified to make the most of paradigmatic cases then used as a reference for the solution of real problems.

For example, regarding the mammography scenario, this can be extended starting from a specific pre-analyzed case in order to define a general reference procedure: what happens if we have to study a mammogram case starting from an actual arbitrary image of a patient? This gives a general procedure that acts as a pattern defined using the algorithm ontology. Specific implementations of algorithms can then be selected based on the properties of the individual media. For example, the general class of FilterAlgorithms has many separate implementations that have equivalent effect but operate on different parameters. Once defined, this specific procedure can then be recorded and stored with the media to provide information for future analysis tasks.

The proposed algorithm ontology supports both types of interoperability. Low-level, syntactic details are described in the Input and Precondition classes while high-level, semantic details are contained in the Effect class. This allows the selection, combination and application of algorithms based on both basic format requirements and on the user desired outcome. This support is useful for both annotation tasks (applying semantically equivalent algorithms to ensure consistency across media objects) and for querying of media based on a description of the processes that have been applied to it (e.g., find media that have had normalization of the light level).

5 Future Work and Conclusions

The processes needed to obtain, elaborate and analyze multimedia objects can be classified, defined and described through a specific ontology of algorithms. This ontology could be used as a base to classify acquired knowledge required in order to solve problems related to the analysis of multimedia objects. The use of the ontology not only will help to solve the problems already known but also similar problems or problems related to analogous contexts.

Development work on the algorithm ontology is ongoing. In particular the questions relating to the quantification and specification of ‘visual’ outcomes from applying an algorithm are challenging. Identifying the main classes of existing algorithms defined and used in literature is extending the ontology. Integration with the 4M infrastructure and the potential use in conjunction with semantic web services and web service choreography technologies is also of interest.

This article has discussed the need for an algorithm ontology in domains such as scientific or medical research, industrial analysis and large-scale digital art analysis. Media, as applied in the scenarios discussed, may undergo multiple analysis, processing and analysis phases. Formalized, structured definitions of media processing algorithms will enable users to classify, identify, locate, apply and record processing and analysis of media throughout its lifecycle.

Acknowledgements

The authors would like to thank members of the W3C Multimedia Semantics Incubator Group for their stimulating discussions about how the Algorithm ontology can be applied to multimedia interoperability. Suzanne Little is supported through a MUSCLE Internal Fellowship (European Commission No. 507752).

References

1. AceMedia: Project, http://www.acemedia.org (2007)
2. Arndt, R., Staab, S., Troncy, R., Hardman, L.: Adding Formal Semantics to MPEG-7: Designing a Well Founded Multimedia Ontology for the Web. Fachbereich Informatik Technical Report Nr. 4. (2007)
3. Asirelli, P., Martinelli, M., Salvetti O.: MUSCLE submission to the Call for a Common Multimedia Ontology Framework Requirements. http://www.acemedia.org/aceMedia/files/multimedia_ontology/cfr/MM-Ontology-Call-Requirements_ISTI.pdf (2006)
4. Asirelli, P., Martinelli, M., Salvetti, O.: Call for a Common Multimedia Ontology Framework Requirements. Harmonization of Multimedia Ontologies Activity. (2006)
5. Asirelli, P. Little, S., Martinelli M., Salvetti, O.: MultiMedia Metadata Management: a Proposal for an Infrastructure. In: SWAP, Semantic Web Technologies and Applications. (2006)
6. Asirelli, P., Martinelli M., Salvetti, O.: An Infrastructure for MultiMedia Metadata Management. In: First International workshop on Semantic Web Annotations for MultiMedia at WWW2006, Edinburgh, Scotland. (2006)
7. Beloozerov, V.N., Gurevich, I.B., Gurevich, N.G., Murashov, D.M., Trusova, Y.O.: Thesaurus for Image Analysis: Basic Version. Pattern Recognition and Image Analysis, Vol. 13(4), (2003) 556-569
8. Bloehdorn, S., Petridis, K., Saathoff, C., Simou, N., Tzouvaras, V., Avrithis, Y., Handschuh, S., Kompatsiaris, I., Staab, S., Strintzis, M.G.: Semantic Annotation of Images and Videos for Multimedia Analysis. In: ESWC 2005, 2^{nd} European Semantic Web Conference, Heraklion, Greece (2005)
9. Boll, S., Klas, W., Sheth A.: Overview on Using Metadata to Manage Multimedia Data. McGraw Hill. (1998)
10. Dasiopoulou, S., Papastathis, V.K., Mezaris, V., Kompatsiaris, I., Strintzis, M.G.: An Ontology Framework For Knowledge-Assisted Semantic Video Analysis and Annotation. In: SemAnnot 2004, 4^{th} International Workshop on Knowledge Markup and Semantic Annotation at the 3^{rd} International Semantic Web Conference. (2004)

11. Dasiopoulou, S., Papastathis, V.K., Mezaris, V., Kompatsiaris, I., Strintzis, M.G.: An Ontology Framework For Knowledge-Assisted Semantic Video Analysis and Annotation. In: SemAnnot 2004, 4^{th} International Workshop on Knowledge Markup and Semantic Annotation at the 3^{rd} International Semantic Web Conference. (2004)
12. DIG35: I3A DIG35 Initiative Group. http://www.i3a.org/i_dig35.html (2001)
13. Dublin Core Metadata Element Set
Version 1.1, http://dublincore.org/documents/1999/07/02/dces/
14. Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: Sweetening Ontologies with DOLCE. In: EKAW2002. (2002)
15. Hollink, L., Little S., Hunter, J.: Evaluating the Application of Semantic Inferencing Rules to Image Annotation. In: KCAP05, the 3^{rd} International Conference on Knowledge Capture. (2005)
16. Hunter, J.: Adding Multimedia to the Semantic Web - Building and Applying MPEG-7 Ontology. Multimedia Content and the Semantic Web: Standards, and Tools, Giorgos Stamou and Stefanos Kollias (Editors), Wiley. (2005)
17. Hunter, J. and Little, S.: A Framework to Enable the Semantic Inferencing and Querying of Multimedia Content. International Journal of Web Engineering and Technology – Special Issue on the Semantic Web (2004)
18. Liberman, L., Menell, J.H.: Breast imaging reporting and data system (BI-RADS). Radiol Clin North Am, Vol. 40, May, (2002) 409-30
19. Little, S., Salvetti, O., Perner, P.: Semantic Annotation of Images. In: A.K.H. Tung, Qiuming Zhu, N. Ramakrishnan, O.R. Zaane, Yong Shi, Chr.W. Clifton, Xindong Wu (Eds.), Proceedings IEEE ICDM Workshops. (2007) 45-51
20. LOM: IEEE Learning Technology Standards Committee's Learning Object Meta-data Working Group, http://ltsc.ieee.org/wg12/ (2002)
21. MARC: Standard, http://www.loc.gov/marc/ (1999)
22. MPEG-7: Multimedia Content Description Interface. ISO 15938. (2003)
23. MPEG-21: http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm (2002)
24. MUSCLE: An Overview of Multimedia Metadata Standards (MUSCLE NoE internal publication), http://muscle.isti.cnr.it/Standards/index.xml (2005)
25. Perner, P.: Data Mining on Multimedia Data. LectureNotes in Computer Science, Springer. (2002)
26. Stamou, G., van Ossenbruggen, J., Pan, J.Z., Schreiber, G.: Multimedia Annotations on the Semantic Web. IEEE MultiMedia, Vol. 13, (2006) 86-90
27. Tsinaraki, C., Polydoros, P., Kazasis, F., Christodoulakis, S.: Ontology-based Semantic Indexing for MPEG-7 and TV-Anytime Audiovisual Content. Special issue of Multimedia Tools and Application Journal on Video Segmentation for Semantic Annotation and Transcoding, Vol. 26, (2005) 299-325
28. VRA Core 4: http://www.vraweb.org/projects/vracore4/index.htm (1999)