Towards a Semantic Content Description for the Visual Arts

Robert David
Institute of Software Technology and Interactive Systems
Vienna University of Technology
Karlsplatz 13, 1040 Vienna, Austria
r.david@semantic-web.at

ABSTRACT

Iconography is the domain of understanding the meaning of historical visual artworks. A formalization of iconographic knowledge can provide a basis for a semi-automatic description of what the content shown on a historical image means without the need for a domain expert. Semantic Web standards can be applied for an iconographic knowledge representation using multiple levels of expressiveness to model the description. This formalized knowledge can then be leveraged to infer the meaning of images based on a description of the depicted visual elements. Having semantically described the content, this information can be used to compare images based on their meaning and find related images with a similar semantic description. We present a novel approach for representing iconographic knowledge using a multi-level semantic model in combination with an iconographic inference rule set. This can be used as a basis for similarity search based on the meaning of the content of historical images.

CCS Concepts

•**Computing methodologies** → **Knowledge representation and reasoning;** •**Applied computing** → **Fine arts;** •**Information systems** → *Clustering and classification;* •**Theory of computation** → *Automated reasoning;*

General Terms

DESIGN, EXPERIMENTATION

1. INTRODUCTION

Iconography studies the meaning of what the content of an image depicts, which is distinct from the artistic style. This meaning is represented by using symbols and allegories. Symbols are objects that mean something on a conceptual level in addition to its literal meaning. Allegories stand for more complex concepts and are usually represented by a combination of multiple symbolic objects. The meaning of the content is dependent on the cultural context of the artwork's creation. The symbolisms used to represent it can be highly complex and difficult to understand. Even if no complex concepts are involved, knowledge about the meaning may not be commonly known anymore. To understand what is represented by the content of a historical image, the viewer has to know about iconography. Even for art historians understanding and interpreting the meaning of an image is a complex task. Art historian Erwin Panofsky defined a methodology [5] for analysing different levels of meaning to help in the analysis of the content. Formalizing some of the simpler aspects of art-historic analyses can support users in understanding the conceptual meaning of historical images.

Even when using an established iconographic system like Iconclass [7], it is very difficult to classify artworks without domain knowledge. Using keyword search for metadata in such a system may still result in several matches, from which the correct one has to be identified. This semantic gap from simple keywords to iconographic concepts has to be bridged automatically in order to reduce the domain knowledge needed for a correct classification. This would provide non-experts with the possibility to reach a higher level of art-historical understanding for an image. It would also be helpful for domain experts as a research tool in the art-historical domain.

Furthermore, a formal iconographic description can also be used for finding similar images based on the meaning of their content. This can provide navigation options for exploratory search and it can be used for the automatic creation of collections based on a theme. Use-cases can be found in research, education and tourism.

We propose an approach for representing iconographic domain knowledge as a formal model including an inference rule set to semi-automatically create an iconographic description for historical imagery. The model also supports the comparison of images based on the content and can be used to find similar images based on their meaning.

2. SEMANTIC CONTENT DESCRIPTION

The model for the Semantic Content Description for image content consists of three parts. First, we define a data model as a knowledge representation for describing the content. The model consists of multiple levels with different semantic expressive power to represent different aspects of the content description. Then a rule set is defined that represents iconographic domain knowledge and can be used by a reasoning system to infer additional semantic descriptions. The last part is the application of similarity algorithms that

work on the semantic descriptions to find images with similar content.

2.1 A model for knowledge representation

As described above the gap between simple keywords and a description using an iconographic classification system is too wide to allow others than domain experts to correctly classify images based on the content. Therefore, we introduce a three level model with increasing semantic expressive power as a knowledge representation for the semantics of image content. We use Semantic Web standards for this representation [2]. They see a widespread use in the Cultural Heritage domain and allow for a flexible semantic data description [3].

Physical level. The first level, the physical level, represents physical objects, including plants, animals and persons, that can be seen on the image and that can be used to describe the content in a direct visual way. No domain knowledge is needed for such a description. We take a painting of the Last Supper as an example. If we describe it at the physical level, the artwork is seen as an image depicting 13 men around a table. No cultural knowledge is needed to describe it that way and we do not have any connections to a religious story.

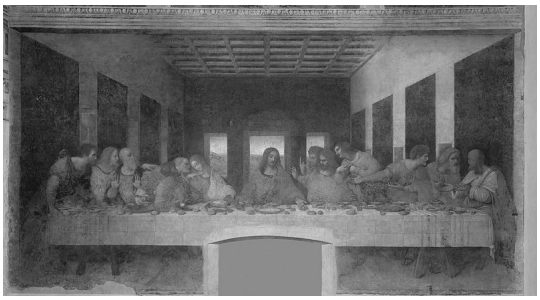

Figure 1: The Last Supper described by three levels: 13 men and a table (physical level), Christianity and religion (conceptual level), and the Last Supper (iconographic level)

Conceptual level. The second level, the conceptual level, represents immaterial concepts like religion, seasons or emotions. These cannot be represented directly in the image, but they can have a symbolic representation by an object or an allegorical representation as a combination of several objects. In contrast to the physical level, a description based on the conceptual level is a semantic description. Here domain knowledge is needed for knowing about the symbolism. With this knowledge we can understand that the painting showing 13 men around a table represents the Last Supper. On the conceptual level the meaning is represented by the concept of Christianity, or more general, that of a religious concept.

Iconographic level. The third level, the iconographic level, represents iconographic concepts and has the highest expressive power. Each iconographic scene is represented by a classification concept. This level is intended to be used by domain experts for a direct iconographic classification. On this level the Last Supper is represented as exactly one concept for the story's scene that the content depicts.

2.2 Taxonomic representation

For each of the three levels we use a hierarchical taxonomy as a representation. For the physical and the conceptual level we use the Art & Architecture Thesaurus AAT [6]. It is defined using the Simple Knowledge Organization System SKOS [1]. AAT contains concepts for many physical things that can be used for representing the physical level. Describing this level using the concepts of a taxonomy is more expressive than using keywords. Even though this level represents no meaning, a controlled vocabulary is needed for defining the rules in the next part. The AAT also contains abstract concepts in the Associated Concepts Facet, that can be used to represent the symbolisms of the conceptual level. For the iconographic level we use the Iconclass classification system [7], which also is available as a SKOS Thesaurus. Having a hierarchy of classifying concepts on each level will be important for the next two parts.

2.3 A rule set for iconographic reasoning

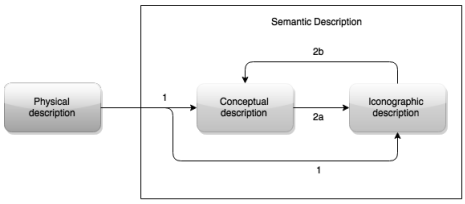

Figure 2: Three level reasoning

For implicit use of iconographic knowledge we propose an iconographic rule set that contains entailments based on the asserted descriptions of the images. These rules can be used by a reasoning system to infer additional describing concepts, as shown in figure 2. Basically, two types of rules can be identified. The first type creates a semantic description by inferring it from a physical description (1). This entails descriptions on the conceptual and iconographic level. The second type improves an existing semantic description by inferring additional concepts on the conceptual and iconographic level (2a, 2b). Both types of rules can be seen as lifting (1, 2a) and lowering (2b) between the three description levels. For improving the rule matching, the hierarchic structure of the taxonomies can be used to reduce the precision needed for matching atoms of the rules. On each rule atom we match not only the concept itself, but the whole subtree below this concept in the hierarchic taxonomy of a level. As some symbols may be interpreted differently based on the cultural context we can add spatial and temporal constraints to the rules. These rules only apply to images that were created in a specific region or during a specific period. Furthermore, specifying the cardinality for rule atoms is important for defining inferences based on the number of objects depicted in the image. As an example, we present rules that can be applied to semantically describe the Last Supper. Starting with a description on the physical level

the content is described by 13 men and a table. As such a scene is likely to represent the Last Supper, we create a rule that infers the Iconclass concept for the Last Supper on the iconographic level based on the physical description of men with cardinality 13 and a table. This inferred description then further infers the concepts for Christianity and religion on the second level based on the Iconclass concept for the Last Supper. The result is a description on each of the three levels.

2.4 Finding images with similar meaning

2.4.1 Taxonomic similarity

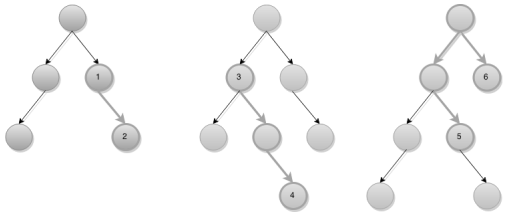

Figure 3: Similarity of concepts on each level: comparing 1 to 2, 3 to 4 and 5 to 6

For comparing the content of an image we compare the descriptions on each level and compute a similarity. The three levels are represented by SKOS Thesaurii and we can use the taxonomic structure as a basis for this computation. For pairs of concepts we can use taxonomic similarity algorithms like Wu & Palmer [4]. As we can describe each of the levels with multiple concepts we have to expand our taxonomic comparison to sets of concepts. Figure 3 shows the three levels and comparisons of concepts therein. We assume two images A and B to compare. Image A is classified with the concepts 1,3 and 5. Image B is classfied with the concepts 2,4 and 6. We compare the two images by comparing their classifying concepts on each level. On the physical level, the concepts 1 and 2 are compared, on the conceptual level, the concepts 3 and 4 are compared and on the iconographic level, the concepts 5 and 6 are compared. The result is a similarity value for the two images on each of the three levels.

2.4.2 Inter-taxonomic similarity

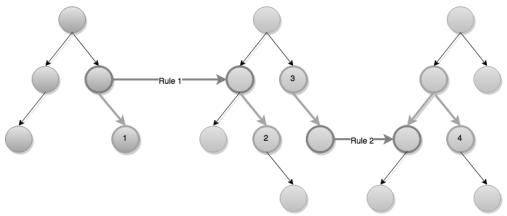

Figure 4: Similarity of concepts on all levels: comparing 1 to 2 and 3 to 4

The disadvantage of using taxonomic algorithms is that a comparison of concepts not of the same level is not possible because they do not origin from the same taxonomy. We have to define an algorithm for computing this inter-taxonomic similarity for concepts that are from different levels. An approach for a computation is a combination of the lifting and lowering rules of the iconographic rule set and the taxonomic reasoning. For comparing two concepts, we find the nearest lifted or lowered two concepts in each of the two taxonomies that are used as atoms in the same rule. These concepts and the connecting rule can be interpreted as a relation for computing the similarity. Using this relation, distance measures can be applied for an inter-taxonomic comparison of the two concepts. Figure 4 shows comparisons of concepts on all three levels. Again, we assume two images A and B to compare. Image A is classified with the concepts 1 and 3. Image B is classfied with the concepts 2 and 4. We compare the two images by comparing their classifying concepts on all levels. We compare the concepts 1 and 2, where concept 1 is on the physical level and concept 2 is on the conceptual level. Rule 1 is interpreted as an edge connecting the physical and the conceptual level. Using this connection a distance value can be computed for the concepts 1 and 2. Then we compare the concepts 3 and 4, where concept 3 is on the conceptual level and concept 4 is on the iconographic level. Rule 2 connects the conceptual and the iconographic level and is used for the computation of the similarity of the concepts 3 and 4. The result are similarity values for the two images spanning all three levels.

3. CONCLUSIONS

We defined a Semantic Content Description for representing iconographic knowledge as a three level model in combination with an iconographic inference rule set. It can be used to infer a semantic description based on what is depicted in an image and to further improve descriptions on the semantic levels. The result is a formal representation of the meaning of the content of images. It can be used in combination with similarity algorithms appropriate for the three level model to find images with a similar meaning. The current state of the work defined the data model and the reasoning system and implemented it as a semantic web application. The next step is the definition of an appropriate iconographic rule set for the reasoning system which has to be done in cooperation with iconographic domain experts. Then a representative and significantly large data base with images and descriptions is needed. Following is the implementation and evaluation of the similarity algorithms in combination with the automatically inferred semantic descriptions. Evaluation of the similarity search will be done using comparisons of keyword-based retrieval, retrieval on each of the three levels and retrieval using the inter-taxonomic approach.

4. REFERENCES

[1] S. B. Alistair Miles. Skos simple knowledge organization system reference. http://www.w3.org/TR/skos-reference/, 2009. [Online; accessed 31-March-2015].

[2] T. Berners-Lee, J. Hendler, O. Lassila, et al. The semantic web. *Scientific american*, 284(5):28–37, 2001.

[3] M. Doerr, S. Gradmann, S. Hennicke, A. Isaac, C. Meghini, and H. van de Sompel. The europeana data model (EDM). In *WORLD LIBRARY AND*

INFORMATION CONGRESS: 76TH IFLA GENERAL CONFERENCE AND ASSEMBLY 10-15 August 2010, Gothenburg, Sweden, Aug. 2010.

[4] R. Mihalcea, C. Corley, and C. Strapparava. Corpus-based and knowledge-based measures of text semantic similarity. In *AAAI*, volume 6, pages 775–780, 2006.

[5] E. Panofsky. *Studies in Iconology: Humanistic Themes in the Art of the Renaissance*. Harper & Row, 1972.

[6] T. Petersen. *Art & architecture thesaurus*. Oxford University Press, 1990. http://www.getty.edu/research/tools/vocabulary/aat/.

[7] H. Van de Waal. *Iconclass: an iconographic classification system*, volume 2. North-Holland, 1974.