Ontology Based Strategies for Supporting Communication within Social Networks

Ivan Kopeček, Radek Ošlejšek, and Jaromír Plhák

Faculty of Informatics, Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic
{kopecek,oslejsek,xplhak}@fi.muni.cz

Abstract. In this paper, ontology based dialogue strategies are presented in connection with the concept of communicative images. Communicative images are graphical objects integrated with a dialogue interface and linked to an associated knowledge database which stores the semantics of the objects depicted. The relevant pieces of information can be linked to the external knowledge distributed in a social network. Exploiting a formal ontology approach facilitates the process of deriving information from relevant texts that can be found in the social network and it simultaneously forms a suitable framework for supporting dialogue communication in natural language. This approach is discussed and illustrated with various examples in this paper.

1 Introduction

Current technologies enable us to associate many useful pieces of information with images that can be used to support users in their retrieval of relevant and interesting information from pictures. Some data are stored automatically when the image is made, typically the date and time of the snapshot and its geographic coordinates, while other data can be retrieved subsequently through the use of various image recognition strategies, social interaction (e.g. people identification in social networks) or manual annotation (e.g. assigning a list of keywords to images within various photo organizers). Regardless of the way the information is obtained it can subsequently be used to support dialogue-based interaction. We refer to images with the ability to communicate with the user in natural language as communicative images.

The basic principles of communicative images have been proposed in [1]. A single communicative image consists of three data structures: (a) graphical content, (b) localization marks that determine the approximate location of the depicted objects and (c) semantic data, as shown in Figure 1. Our approach exploits the Scalable Vector Graphics (SVG) format [2] to encode all these data structures in a single file. The semantics are encoded as Web Ontology Language (OWL) ontologies [3] which provide a suitable formalism for information retrieval and machine-generated dialogues.

For example, let us imagine a holiday photo in JPEG format, as shown in Figure 1: My wife Jane in front of a castle. To make this photo communicative, it has to be converted into SVG format and stored with the original JPEG data. The supporting OWL ontology that is linked from the SVG, should define terms like “castle” and “wife”. The annotation data, which is also embedded in the SVG file, links specific

P. Sojka et al. (Eds.): TSD 2014, LNAI 8655, pp. 571–578, 2014.

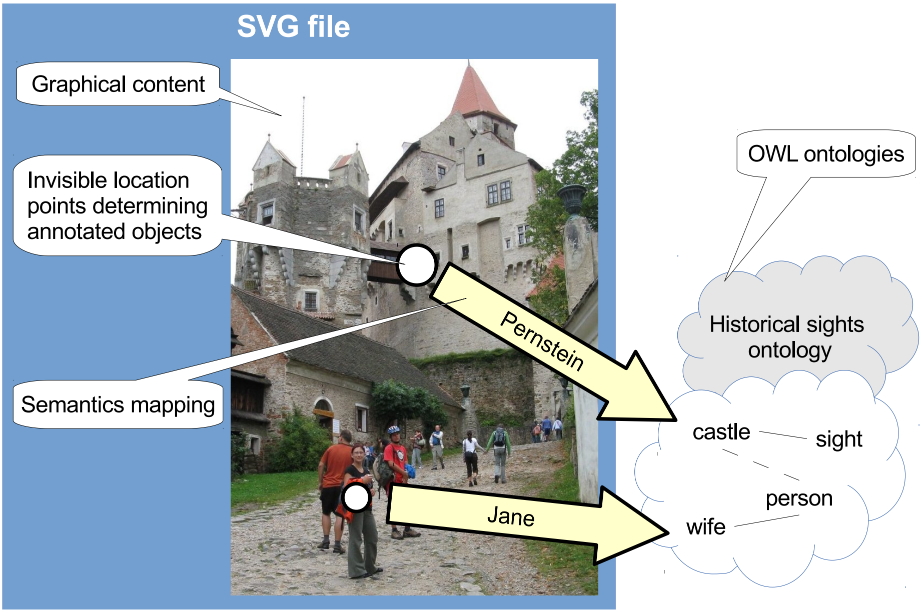

Fig. 1. Structure of a communicative image

graphical elements with terms in the ontology. Provided that the annotation data include the information that the wife has the name Jane and the castle is Pernštejn, the image is able to answer questions like "What is the castle behind Jane?". Moreover, connecting to a general ontology of historical sights, the communicative image would be able to provide the user with details about the castle.

Dialogue-based communication is essentially interactive. It enables us to control the complexity of the information and it prevents the user from being overburdened by superficial information. However, the ability of an image to communicate informatively depends on the ability to associate sufficient relevant information with the image. But making picture annotations and creating ontologies by hand is very labor-intensive and tedious. Furthermore, gathering such data by means of existing image recognition and auto detection techniques cannot always solve the problem because they are still far from being able to fully describe an analyzed picture. Our goal is therefore to enable images to actively learn from the communication. New pieces of information are provided either directly by the user or indirectly by the social community being active in a social network.

Consider the holiday photo discussed above and assume that this photo has no semantic data assigned so far. The following fragment of a dialogue demonstrates the ideas and principles of active learning.

Example 1 (*U = User, P = Picture, // comment*):

U: What is the castle behind Jane?
// At this moment, the communicative picture can assume that it depicts

some castle and some object called Jane. The picture takes over the initiative to learn more about these two things.
P: Who or what is Jane?
U: Jane is my wife.
// Picture semantics can be extended with the fact that there is a woman, the wife of the user, called Jane.
P: Based on geographic coordinates the castle would be Pernštejn. Is it possible?
// An explore external shared knowledge database covering geographic data of sights is now needed to identify the castle according to its position.
U: Yes, it is Pernštejn.
// The user recalled the trip and confirmed the information. The picture extends its semantics with this fact for later re-use.
U: When was Pernštejn built?
// At this moment, the picture tries to retrieve details about Pernštejn castle, either searching the available ontologies or asking the social community for help.
P: I'm sorry but I don't know.
// No information was found at this moment and then the picture asked the social community for assistance. It takes some time to get response but next time the picture should know the answer.

This dialogue demonstrates a way of overcoming the lack of information when neither the user nor the associated knowledge databases are able to provide valid semantic data. In this case, the user question can be stored in the image and shared with social community. This process has to be carefully managed so as to prevent violating the user's privacy and to protect the community from being overburdened with such questions. To optimize this task, question similarity detection, well-formed question verification and answer credibility need to be taken into account.

2 Related Work

Our approach to communicative images is based on formal modelling through ontologies. Ontologies are considered one of the pillars of the semantic web. Many papers have been published dealing with ontology-based intelligent data retrieval from databases [4], the annotation of graphical data [5,6] and from e-learning [7,8]. In [9], Chai et al. proposed an intelligent photo album enabling collections of family photos to be organized and searched by means of ontologies.

In the context of communicative images, dialogue management has to be able to achieve high efficiency in managing the image ontology and provide suitable dialogue strategies. Some approaches proposed in [10,11,12] that enhance the efficiency of the dialogue manager exploiting the knowledge bases have been presented. The dialogue management is separated from the domain knowledge management. In [13], it has been argued that this separation reduces the complexity of the systems and enhances further extensions.

3 Dialogue Strategies for Communicative Images within Social Networks

Communicative images represent dialogue systems that support providing information about non-textual data and other related modes of communication. The pieces of information that are exploited in this process are of the following types:

- *Type 1*: The information directly integrated with the object at the moment of creation, e.g. date and time the picture was shot, geographic coordinates, technical data of the camera, etc.
- *Type 2*: Basic semantic information about the picture which is incorporated directly into its SVG format. It is possible that there is no such information when the picture is created – it is provided by the users during the process of communicating with the picture.
- *Type 3*: The global information situated in the ontology forms in the cloud. Because this ontology saves all possible pieces of information based on the communication within the network, it can be too large to be integrated with the format of the picture. The structure of the ontology, however, makes finding relevant information or providing basic inference processes feasible.
- *Type 4*: The pieces of information spread across the network and internet in various forms, mainly as text. Such information can be converted into the format of the ontology, which enables more flexible use. But if this conversion is not feasible, it can be presented in its original form.

Based on this typology, the process of supporting communication with the user can be characterized as follows.

- Queries about Type 1 information are answered based on the integrated graphical ontology (see [14]).
- Queries about Type 2 information are answered based on the ontology which is integrated with the format of the picture.
- Queries about Type 3 information are answered based on the ontology in the cloud.
- Queries about Type 4 information are answered based on the ontology which is dynamically created in real time during the communication. If it appears that the conversion of the pieces of information into the ontology is neither feasible nor efficient, they can be presented in their original forms.
- Within the social network, the users can add comments, descriptions and answers to the questions of other users about the picture at any time.
- If the query cannot be understood by the system, the system either provides metacommunication with the user or passes the query to other users of the social network to be answered by them. In such cases, it is assumed that they will not be answered in real time, and it is possible that they will remain unanswered.

4 Testbed Implementation

The concept of ontology-based communication strategies is tested and verified by means of a testbed application, which has been developed to deal with communicative images. Its client-server architecture enables us to develop many specialized clients as well as to share the knowledge on the server side.

Clients, typically web applications or plugins to social networks, handle a user's interactions with the image and redirect these interactions to the remote server.

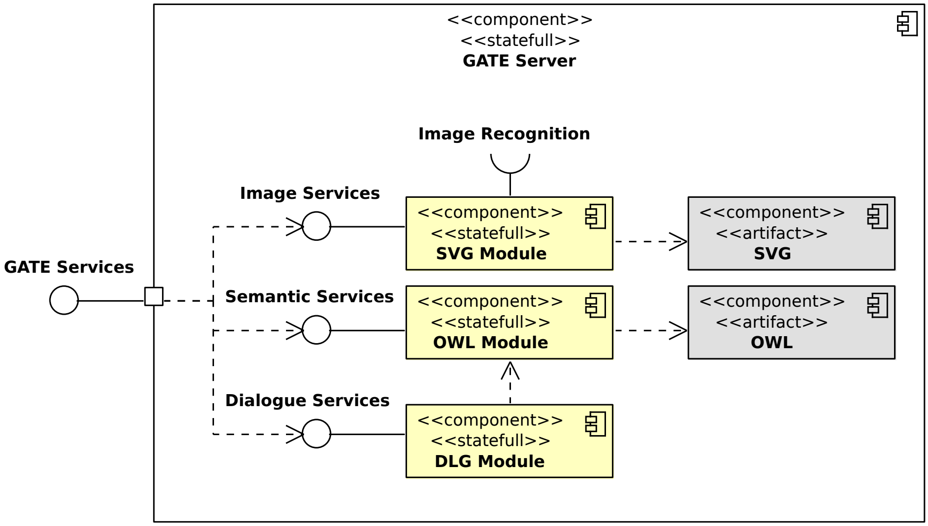

Fig. 2. Component architecture of GATE server

GATE server is designed as a component-based Java enterprise application, as shown in Fig. 2. It provides high-level session-oriented remote services via *SOAP* or *RESTfull* APIs. These services enable clients to upload an image and then to explore it by means of natural language. The server consists of several lower-level components providing services for the graphical content management, semantics management and dialogue subsystem.

The *SVG Module* handles graphical content of a single uploaded image. Provided services enables to explore the SVG Document Object Model (DOM). Either SVG or raster image can be uploaded at the beginning of the communication. Raster images are automatically wrapped with initial SVG content. When connected to external *Image Recognition* services, this module can automatically analyze the image and get initial semantic data by employing domain-specific algorithms, e.g. face recognition [15,16] or similarity search algorithms in large image collections [17,18,19].

The *OWL Module* is responsible for managing the semantic data encoded in OWL format. Annotation data and ontologies encoded in the uploaded communicative image are gathered and processed automatically. Additional ontologies can be added during the communication. Provided services cover ontology management, low-level traversal of the OWL DOM tree, ontology reasoning and information filtering.

Several OWL ontologies have been integrated into this module so far. A *Graphical ontology* [14] can be used to express significant or unusual visual features of the objects depicted as well as their location and mutual position, enabling the annotator to describe the scene in terms like "Abnormally big cat in the upper left corner." or "Oval pool in front of a house." (expressions defined by the ontology are underlined). A *Family ontology* can be used to classify people or to automatically infer their family relationships. A *Sights* ontology provides the vocabulary and background knowledge to describe interests, historical buildings and monuments.

The *DLG Module* represents dialogue subsystem. It is responsible for parsing and understanding questions in natural language and composing answers. This module cooperates closely with the *OWL Module* to analyze the meaning of words.

At present, only a simplified version of this module is implemented. This version supports questions in What-Where Language (WWL) [6], having the format "where is what", "what is where" or "what some object is". Moreover, to design an efficient dialogue management, we analyzed a corpus of relevant user utterances and identified a relatively small fragment of the natural lan- guage which is used by most users. Based on this analysis, we defined the templates that consist of slots specifying the pieces of information to be acquired from the user. These templates are also integrated into the system. For instance, the question "*Who is between me and my wife?*" is resolved using the template "*Who is between SLOT1 and SLOT2?*". The system expects both the *SLOT1* and the *SLOT2* to be filled with the specific individuals from the "*Person*" and "*Object*" classes.

This basic implementation was tested by visually impaired people whose goal was to interact with well-annotated pictures by means of WWL and give us their personal feedback. Although the current functionality is still very limited, there is no social network integration nor any generation of ontologies from texts at the moment. Despite this, the preliminary responses of users are promising and show that the concept of communicative images is feasible and seems to have potential in many application domains.

5 Application Domains

Precisely annotated communicative images are applicable to e-learning systems where the tutor prepares study materials thoroughly by hand. In this case, connecting the depicted objects with a broader knowledge database, e.g. historical, can provide students with a "window" into the whole topic of a field of study. For instance, a well-annotated picture of the battle of Austerlitz enables the students to discuss scene and naturally learn who Napoleon was, when the battle took place and who won.

Communicative images also support the accessibility of graphics for people with special needs. A dialogue with the image held in natural language makes the graphical data accessible especially to visually impaired people. The users are not limited to a simple summary of the image's content. Since the data is structured and related to different parts, objects and aspects of the image, a complex dialogue can be undertaken, ultimately leading to a more natural and fulfilling experience for the users.

Any user can benefit from the ease of access to photo albums. The user can browse family pictures while being reminded of the age and names of the people in the photos, their birthdays, the time and occasion the picture was taken. Moreover, ontologies used in communicative images could support efficient searching for specific images. The search query referring to specific objects or relations depicted in a source image can be specified either in natural language or using an intelligent graphical user interface. Pictures that have satisfy a query can be grouped into thematic albums creating and maintaining some order in large collections of images.

6 Conclusions and Future Work

In this paper, we have outlined basic principles of communicative images as well as the general architecture of the system. Currently, the concept of communication images suffers from the problems associated with the automatic gathering of semantics from available sources. The aim of this paper has been to propose a formal approach of obtaining semantic data from relevant texts. Nevertheless, the concept of communicative images has still many problems yet to be addressed. For instance, there is a gap between semantic models and dialogue strategies. We have to carefully prepare and fine-tune the dialogue subsystem for each concrete domain ontology by hand, instead of generating dialogue strategies automatically from the internal structure of the ontology provided. Also continual enhancement and enlargement of knowledge bases together with automatic learning from dialogues pose a great challenge. In spite of these obstacles, the preliminary results show that our approach is implementable and promises important application to many domains.

References

1. Kopecek, I., Oslejsek, R.: Communicative images. In: Dickmann, L., Volkmann, G., Malaka, R., Boll, S., Krüger, A., Olivier, P. (eds.) SG 2011. LNCS, vol. 6815, pp. 163–173. Springer, Heidelberg (2011)
2. Dahlstóm, E., et al.: Scalable vector graphics (svg) 1.1, 2nd edn. (2011)
3. Lacy, L.W.: OWL: Representing information using the Web Ontology Language. Trafford Publishing, Victoria BC (2005)
4. Muñoz, A., Aguilar, J.: Ontological scheme for intelligent database. In: Proc. of the 11th WSEAS Int. Conf. on Computers, ICCOMP 2007, pp. 1–6 (2007)
5. Schreiber, A.T.G., Dubbeldam, B., Wielemaker, J., Wielinga, B.: Ontology-based photo annotation. IEEE Intelligent Systems 16, 66–74 (2001)
6. Kopeček, I., Ošlejšek, R.: Gate to accessibility of computer graphics. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A.I. (eds.) ICCHP 2008. LNCS, vol. 5105, pp. 295–302. Springer, Heidelberg (2008)
7. Henze, N., Dolog, P., Nejdl, W.: Reasoning and ontologies for personalized e-learning in the semantic web. Educational Technology & Society 7, 82–97 (2004)
8. Lemnitzer, L., et al.: Using a domain-ontology and semantic search in an e-learning environment. In: Innovative Techniques in Instruction Technology, E-learning, E-assessment, and Education, pp. 279–284. Springer, Netherlands (2008)
9. Chai, Y., Xia, T., Zhu, J., Li, H.: Intelligent digital photo management system using ontology and swrl. In: Proc. of the 2010 Int. Conf. on Computational Intelligence and Security, pp. 18–22. IEEE Computer Society (2010)
10. Nyrkko, S., et al.: Ontology-based knowledge in interactive maintenance guide. In: Proc. of the 40th Annual Hawaii Int. Conf. on System Sciences. IEEE Computer Society (2007)

11. Pérez, G., et al.: Integrating owl ontologies with a dialogue manager. Technical report, Spoken Language Technology Workshop 2006, pp. 134–137. IEEE, Phanouriou (2006)
12. Araki, M., Funakura, Y.: Impact of semantic web on the development of spoken dialogue systems. In: Lee, G.G., Mariani, J., Minker, W., Nakamura, S. (eds.) IWSDS 2010. LNCS (LNAI), vol. 6392, pp. 144–149. Springer, Heidelberg (2010)
13. Flycht-Eriksson, A.: Design and Use of Ontologies in Information-providing Dialogue Systems. Dissertation, Linköping studies in science and technology, thesis no. 874, School of Engineering at Linköping University (2004)
14. Ošlejšek, R.: Annotation of pictures by means of graphical ontologies. In: Proc. Int. Conf. on Internet Computing ICOMP, pp. 296–300 (2009)
15. Bartlett, M.S., Movellan, J.R., Sejnowski, T.J.: Face recognition by independent component analysis. IEEE Transactions on Neural Networks 13, 1450–1464 (2002)
16. Haddadnia, J., Ahmadi, M.: N-feature neural network human face recognition. Image and Vision Computing 22, 1071–1082 (2004)
17. Batko, M., Dohnal, V., Novák, D., Sedmidubský, J.: Mufin: A multi-feature indexing network. In: Proceedings of the 2009 Second International Workshop on Similarity Search and Applications, pp. 158–159. IEEE Computer Society (2009)
18. Jaffe, A., Naaman, M., Tassa, T., Davis, M.: Generating summaries and visualization for large collections of geo-referenced photographs. In: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval, pp. 89–98. ACM (2006)
19. Abbasi, R., Chernov, S., Nejdl, W., Paiu, R., Staab, S.: Exploiting flickr tags and groups for finding landmark photos. In: Boughanem, M., Berrut, C., Mothe, J., Soule-Dupuy, C. (eds.) ECIR 2009. LNCS, vol. 5478, pp. 654–661. Springer, Heidelberg (2009)