An Opportunistic Approach to Adding Value to a Photograph Collection

Ashley Smith, Leslie Carr, Wendy Hall
School of Electronics and Computer Science
University of Southampton
{ads04r | lac | wh}@ecs.soton.ac.uk

Abstract

The Semantic Web can, among other things, be used for photograph annotation. Many implementations of this idea exist, but all are limited by the fact that a human must manually create the annotations for the photographs, often using a program with which he or she is not familiar. This poster discusses an opportunistic method of photograph annotation that uses logical inference in conjunction with existing data from various sources in order to obtain information about the images being annotated.

1. Introduction

The Semantic Web, or more specifically its underlying framework RDF, is a useful method of annotating photographs. RDF is a method for storing simple facts about resources (in this case photographs are the resources) in a standard format that is understandable to machines, meaning that searches of large photo collections can be much richer, as typically an image collection can only be searched by date or file name. However, for these annotations to exist they must first be created. The process of annotating photographs manually is a time-consuming process. Several applications have already been created that attempt to address this problem but ultimately their uptake has been far from overwhelming, as complex or novel user interfaces generally have a steep learning curve and most computer users simply do not wish to change the way they work [Sauermann and Schwarz, 2004]. We present a possible solution to the problem by having the computer annotate the images using an unobtrusive and completely automatic process without the need for the user to do anything at all. This may be achieved by evaluating various different information resources, both Semantic Web enabled and otherwise, and applying logical inference to the known facts in order to derive otherwise unknown information about the photos in question. Although not exhaustive or definitive, this provides a small amount of basic information without a tedious manual annotation process.

2. Related Work

The Gnowsis [Sauermann and Schwarz, 2004] is a server/client system. The server element runs as a knowledge store in the background, and the client side actually consists of a collection of plugins for various popular applications such as Microsoft Outlook and Mozilla Firefox. This allows the user to create semantic links between resources of various different types, such as linking a web site to an email or an address book contact if they are related. Its main flaw is its lack of automation. In most instances the user is required to explicitly define links between resources. In contrast, Haystack [Quan et al, 2003] is a full desktop organiser that has an RDF back end but does not actually display any of its underlying data to the user. It is aimed at completely novice users, but it is of no real use to people who already use computers daily and have grown accustomed to using more mainstream software to do their work.

In the context of photo annotation, several methods exist for adding meta data to images. These range from web-based 'gallery' sites to applications that are run on a local computer. Photofinder [Shneiderman and Kang, 2000] is a locally run application that allows the user to add semantic meta data (albeit not RDF) to an image collection. It is a very simple 'drag and drop' style user interface that can be used to 'pin' notes to an image. It allows a certain level of bulk annotation, but is very limited due to its strict database structure. Photobucket[1] is a popular gallery that is used by bloggers and eBay users to store small collections of images. It allows the user to add a title and a description to all images and organise them into collections, but does not allow any form of semantic markup and all images must be annotated individually. Flickr[2] is much more advanced. It allows the user to specify image regions, useful for identifying individuals in group photographs, and allows for bulk annotation. However it has no native RDF output, despite having a very rich API, and it has no way of adding semantic data to photos, limiting annotation to keywords.

1 http://photobucket.com/
2 http://www.flickr.com/

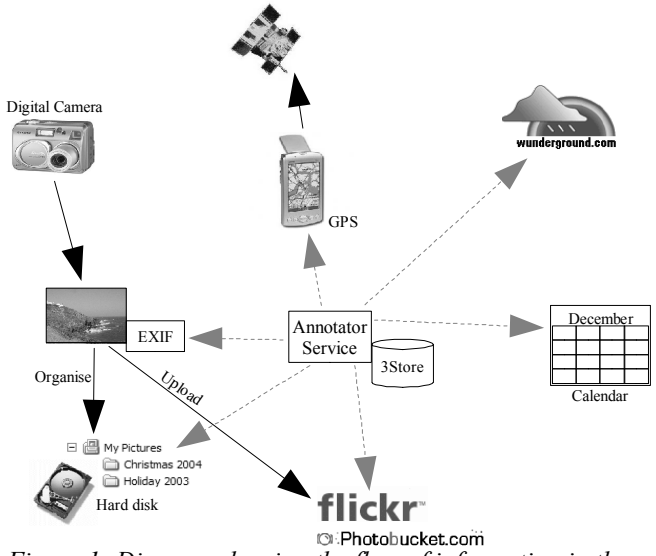

Figure 1: Diagram showing the flow of information in the proposed system.

3. An Opportunistic Approach

The beauty of the Semantic Web is that it allows many different sources of information to be shared in a common format. Supposing a photograph was taken on a certain date by a certain person. Querying that person's calendar may indicate that they were on holiday in Hawaii at the time, and the location recorded by the GPS function on their PDA at around the same time confirms this. We can deduce from this very sparse information that the photo was taken in Hawaii, and depicts a holiday. If we know the time and date a photo was taken, we can also query a local weather station to determine the weather in the area at that time, and apply this to the photograph's annotations as well. If the photographer has an online account at a service such as Flickr or Photobucket, we can query this service to obtain the photograph's status (public or private) and also any manual annotations the photographer may have made.

In addition to the factual data available, we can deduce other slightly less certain information in order to support what we already know. We can employ a weighting system (similar to Google's PageRank[3]) that calculates the probability that a photograph belongs to a particular collection. For example, if two photographs exist and they were both taken by the same camera within five minutes of each other, it is much more likely that they are related in some way than if they were taken hours apart, or at the same time with different cameras. Equally, if two photos were taken and both correspond to an event described as a conference in the photographer's calendar, it is more likely that they are loosely related than if one is a conference and one is vacation time.

Finally, it is reasonable to derive additional annotations from existing RDF triples, which may be distributed. For example, if person A goes on holiday and emails a picture he/she took while there to person B, with no annotations other than the EXIF data, person B should be able to query person A's calendar to determine the location of the photo.

3 http://www.google.com/technology/

4. Conclusions and Future Work

It would be possible and desirable to implement a piece of software that runs as a service, invisible to the user, that keeps an eye on any new photos being added to a repository, be it a folder on a hard disk or an online picture storage account, and queries all relevant information from other sources, in an attempt to gain enough information about the image's context to annotate it to some level of completeness (Figure 1). The service would interact with a knowledge store, which can be queried, and would require no user interface as the user does no work. A separate application could be used to interact with the knowledge store and add additional, manual annotations should they be required, as well as query and manage the annotations that already exist. As the knowledge store will use an open query language such as RDQL, other applications and plugins for existing applications may be written, so that the user may choose an user interface with which he/she is comfortable.

We propose a study into non-content-based similarity of images. We intend to carry out a diary study accompanied by a series of interviews in order to determine how people store their digital images, and attempt to identify common habits. We intend to use the results to determine the feasibility of annotating photographs without actually looking at them, but knowing rudimentary information about their context, such as the date they were taken, the geographical location that they were taken and in which directory the photographer has chosen to store them. Links to external but relevant information, such as the user's calendar, a list of worldwide events, and news and weather feeds for the relevant time period would be of benefit as well. This direction also has the potential to collaborate with other research areas, such as image processing, in order to determine the physical similarity of two images, rather than just their contextual similarity.

References

[Quan et al, 2003] Dennis Quan, David Huynh, and David R. Karger. Haystack: A Platform for Authoring End User Semantic Web Applications. In proc. 2nd International Semantic Web Conference (ISWC2003). 20-23 October 2003, Sanibel Island, Florida, USA.

[Sauermann and Schwarz, 2004] Leo Sauermann and Sven Schwarz. Introducing the Gnowsis Semantic Desktop. In proc. 3rd International Semantic Web Conference (ISWC2004). 7-11 November 2004, Hiroshima, Japan.

[Shneiderman and Kang, 2000] Ben Shneiderman and Hyunmo Kang. Direct Annotation: A Drag-and-Drop Strategy for Labeling Photos. In proc. International Conference Information Visualisation (IV2000). 19-21 July 2000, London, England.