Capturing Diverse Usage Contexts for Web-based Images

Randi Karlsen
Computer science department
University of Tromsø, Norway
randi@cs.uit.no

Joan Nordbotten
Department of Information and Media Science
University of Bergen, Norway
joan@infomedia.uib.no

Abstract

The rapidly increasing number of image collections on the Internet require improvements in our image retrieval systems if these collections are to be easily available for browsing and searching by information seekers. Current techniques for image retrieval have known shortcomings that make it difficult to search for images based on their semantic content. We propose that an increased use of image context information can improve identification of image semantics, and may thus contribute to closing the gap between user needs for semantic image retrieval and the capabilities of current image retrieval systems. In this paper we present a new category of image context, called Usage Context*, describe how Usage Context from multiple sources can be combined, and show how Usage Context information can be specified using a new context descriptor vocabulary CTXT.*

1 Introduction

In recent years image collections available on the Internet have increased both in number and size, and represent a huge amount of important information. Both privately and publicly owned collections of images are available for browsing and searching, and the number of users and application areas are increasing. To avoid that numerous images become hidden treasures, it is crucial to improve current techniques for image description, location and retrieval.

Information seekers typically describe their information need in terms of the semantic content they want in the image(s) retrieved. However, many image retrieval systems have only rudimentary ability to support semantic image retrieval [1]. We believe that an extended use of image context information can enhance automatic identification of image semantics and thus may contribute to closing the semantic gap between user needs and current capabilities of image retrieval systems.

There are multiple forms of image context information, ranging from the date/time and place where the image was created to the context(s) of its use. This paper defines *Usage Context* and focuses on representation of Usage Context information for images. Usage Context describes the environment in which an image is used, typically an image collection or an illustrated document, and is attached to images for the purpose of identifying semantic content of images. As images are shared and distributed, copies of an image may be used in a number of environments. Combining Usage Context information from multiple sources can be useful for improving determination of the content of an image, as well as for increasing understanding of, and possible uses for an image.

In this paper we investigate how multiple Usage Contexts can be combined to provide enhanced knowledge of the semantics of individual images. Thereafter, we propose a vocabulary, CTXT, for context specification and demonstrate how Usage Context information can be represented so that the characteristics of each context type are sustained and made available for use in image retrieval.

The remainder of the paper is structured as follows. Section II presents a background on image retrieval, current use of image context information, and introduces Usage Context. Section III describes different types of Usage Context for images, while Section IV describes combined use of the context information. In Section V we present the basic version of the context vocabulary CTXT, and Section VI summarizes this proposal.

2 Background

2.1 Image retrieval

Before retrieving images from the Web, there are several tasks that need to be carried out; including data gathering, image analyzing, and indexing [2]. During data gathering, a crawler traverses the Web collecting images and related data. This information is then analyzed to compute image features that are used for retrieval and indexing. Depending on where the image features are extracted from, there

are two main image retrieval approaches; text-based and content-based retrieval.

In *Text-based image retrieval* (TBIR) image features are extracted from image annotations that may describe semantics of objects in the image as well as the situation or activities depicted in the image [3]. Unfortunately, annotations may be biased, since they represent the annotator's understanding of the image and are described in his/her vocabulary. In addition, annotating images is primarily a manual, time consuming task, which makes it likely that an image is annotated with only one or a small subset of its possible semantic interpretations.

Content-based image retrieval (CBIR) [4] has been developed to support image retrieval in areas where textual descriptions of images are either not possible or impractical (such as satellite images, finger prints, and x-rays). The underlying idea of CBIR is image matching, where the structural characteristics of an input image are matched to the images in the database. CBIR supports automatic registration of low-level image features, such as color, shape or texture, but lacks the support for image retrieval based on high-level semantic concepts. This semantic gap [1], which represents the mismatch between semantic user requests and the capabilities of current CBIR systems, is a major problem for general CBIR systems.

2.2 Image Context

The shortcomings of traditional image retrieval techniques have in resent years lead to research activities on including context as an additional source of image metadata. This section briefly summarizes some of these approaches.

A commonly used definition of context is described in [5]: *"Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves"*.

Dey [5] notes further that context characteristics for entities can include location, time, activity, and identity. These four characteristics are, according to Dey, all primary context types, that can be used as indices to find secondary context types.

When the entity-type is an image, the *time* and *location* characteristics are currently the most used context information. Time and location (such as GPS coordinates) have been used for clustering images into collections [6, 7], for determining relationships between the image and real world objects [8], location-based image search [9], for navigating through an image collection [10], and for detecting information on weather, local time, daylight status and season [11]. A number of systems combine image location information with maps, see for instance Flickr [1], Smugmug [2] and Woophy [3]. *Identity* of the photographer together with social context is in [12] used for automatically resolving the identity of subjects in an image.

We find that context can characterize two different types of situations for an image, i.e. *image creation* (or capture) and *image usage*. Our description of context reflects this distinction, and we identify two context types for images; Usage Context and Creation Context.

- *Usage Context* represents information about the environment in which an image is used, for instance an image collection or a document where the image is used as illustration. Context information can be a textual description of a collection, a document abstract, keywords describing the collection or document, or text surrounding an image.

- *Creation Context* represents information about the environment in which the image was taken, and may include sensor data, information about an area or objects, the purpose of the image, and more.

Our current focus is on Usage Context and we investigate how context information from multiple usage environments (i.e. collections and/or documents) can be used to infer semantic content of images.

Usage Context has to some extent been used in previous work. For an image included in a document or web page, keywords can be automatically extracted from surrounding text, image caption, page title, and image file name. This approach is followed in research such as [13, 14] and in systems such as Google Image Search [4]. In [15, 16, 17] collection metadata is used for selecting the most relevant text document collections for distributed information retrieval.

Our work differs from previous in that we focus on identifying, capturing, and combining Usage Context information from different sources. Our goal is to let Usage Context from multiple environments be combined and collectively serve as indicators of the semantic content of the image.

2.3 Image metadata representations

Metadata, representing information about the image, is recorded for subsequent use in image retrieval. There are different metadata formats available, representing image metadata somewhat differently. The Exif (Exchangeable image file) format[5], which is used by most digital cameras, includes a number of metadata tags holding information

[1] http://www.flickr.com/
[2] http://maps.smugmug.com
[3] http://www.woophy.com/map
[4] http://images.google.com
[5] http://exif.org/

such as camera settings, date and time, location, as well as user provided textual information. Some metadata, such as date/time and camera setting, are automatically provided by the camera. Other types of metadata are, if available, provided from other sources, for instance GPS units or manual annotation.

Other important standards for metadata specification include Dublin Core[6], MPEG-7[7], CIDOC/CRM[8] and RDF[9]. The first 3 of these provide standard attribute sets for describing characteristics of multiple media objects. RDF (Resource Description Framework) is an XML-based framework for representing these attributes in a standardized way. The ability of RDF to represent image metadata is demonstrated in [18] and in [19] where EXIF data is transformed to RDF.

3 Usage Context for images

Collections and documents are two important types of environments that are sources for Usage Context information. This section describes these context sources, and explains how information from the sources can be understood.

3.1 Image collections

An image collection is a set of images that has one or more properties in common. Typically, the images in the collection are brought together for some specific purpose, such as recording i) a specific event, e.g. a holiday or anniversary, ii) a time period or, iii) thematic content [20].

An image collection can be generated either automatically or manually, where *automatic* generation is based on available image metadata, such as date and time information [6]. With *manual* generation of collections, users cluster their images according to some criteria that might not be represented in the image metadata. A thematic collection based on semantic knowledge of the images, is typically manually generated.

Collection-level metadata include information i) derived through computer analysis of the content of the collection, and ii) supplied by the collection provider, typically through annotations [15]. The last type of metadata, which may include collection title, application area, purpose, and topic descriptions, are of specific interest in this paper as it provides additional information compared to what is available from each image alone, and may thus indicate semantic interpretation of the images within the collection.

We are in this work specifically interested in collection information containing semantic information, and are thus relying on the willingness of the user to annotate his or her collections. A collection can include a large number of images, and we can not assume that a user will annotate each image. However, annotating the collection itself is far less time consuming, and we therefore find it reasonable to assume that users may provide collection descriptions.

A *collection description* is part of the *collection-level metadata*, and may include collection title, scope, purpose of the collection, what kind of images are included, a topic description for the collection, and other information. This description includes semantic information about the collection, which we will use to indicate a semantic interpretation of the images within the collection.

3.2 Documents

An illustrated document represents another type of image context. Metadata describing a document will, in the same way as collection metadata, represent Usage Context information for the images used in the document. A common characteristic of images in a specific document is that they are used as illustrations to support some part of the content of the document. Document information useful as Usage Context information includes; i) general descriptions of document content, such as an abstract or keywords and ii) text surrounding an image.

A general document description provides information about the content of the document, and describes Usage Context for all images in the document. Surrounding text is, in contrast to document descriptions, specific to the particular image that it surrounds, and may therefore give a more direct indication to what the content of the image is.

3.3 Usage Context information

Usage Context sources provide different types of Usage Context information, that carry somewhat different semantic meaning as described below.

- A *collection description* or *document abstract* gives a general description of the collection/document. This does not represent a direct description of the content of any of the images in the collection/document, but can provide an indication of what an image is about.
- Documents and collections can also be described through *keywords*, which capture the essence of the collection/document in a few, carefully chosen, terms. Compared to descriptions or abstracts, keywords may provide more precise information about the topic of the collection/document.
- *Surrounding text* represents a context which is local to a specific image. While a document description provides general context information that applies to a set

[6] http://www.dublincore.com
[7] http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm
[8] http://cidoc.ics.forth.gr/
[9] http://www.w3.org/RDF/

of images, surrounding text provides context information specific to a particular image.

The above description illustrates that we can attach different interpretations to the different sources of Usage Context information. It is therefore important that the representation of image Usage Context data distinguishes between the different sources.

4 Combining usage context information

An image may be useful in different situations and may therefore be included in a number of collections and/or documents, each with its associated usage context information. Our goal is to combine Usage Context information from different sources to better determine the content, understanding of, and possible use of an image.

4.1 From usage context to semantic image metadata

Usage Context information, such as a collection description or a document abstract, represents a semantic description of the collection (or document) where a number of images are included. If an image is included in several collections, there may be equally many collection descriptions related to the image, each representing a separate Usage Context description.

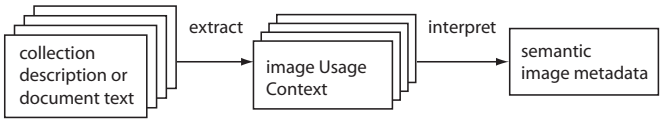

Figure 1. From usage context to semantic image metadata

Figure 1 shows the process of transforming information from different context sources into semantic metadata for an image. Usage Context information can be automatically extracted from collection/document metadata and document content, and stored together with the image using the CTXT vocabulary presented in Section V. The CTXT vocabulary is designed for structuring Usage Context information so that the semantic meaning is sustained and reasoning about the different types of context information is possible.

The extracted information may include text and/or keywords from a collection description, document abstract, title, or surrounding text, together with a collection/document identifier, document type, purpose of the collection, and other types of information. This Usage Context information can subsequently be used as a basis for obtaining a semantic understanding of the image. Since Usage Context provides information about the environment where an image is used, and is not a direct description of the image, Usage Context data is used as an *indication* to what the image is about.

Interpretation of Usage Contexts can be an automatic process, controlled by rules determining interpretation policies. One approach is to compare Usage Context from different sources (for instance keywords from collection descriptions and document abstracts) to identify topically overlapping contexts that collectively support a semantic view of an image, or identify topically disparate context that may indicate different semantic views of an image. Rules can be used to determine the relative importance of context information, how information is compared, how ontologies are used, and how information is combined to infer image semantics.

To illustrate how a combination of Usage Contexts can be used, assume two image collections; one with images of "Boats", $Coll_{Boats}$, the other with images on the topic "Fishing", $Coll_{Fishing}$. An image appearing in both collections is most probably of a fishing boat. We also assume a document, describing the boat type "Smack", which contains a number of images used for illustration. In this example we represent all images in the document as an image collection, $Coll_{Smack}$, see figure 2. The document contains different types of images, showing for instance boats, equipment, boat-houses or activities. However, if an image is included in both $Coll_{Boats}$ and $Coll_{Smack}$, we have reason to believe that the image depicts a boat of type smack.

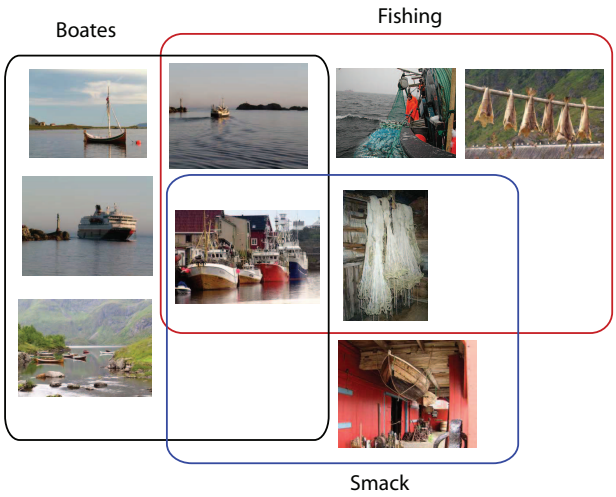

Figure 2. Image collections

We may also deduce other types of information based on these collections. The fact that images from the "Smack" document have been used as illustrations provides useful information. Assume we are writing a paper and need an image of a boat for illustration purposes. Among the images in $Coll_{Boats}$ some are suited as illustration, while oth-

ers are not. If an image is included in both $Coll_{Boats}$ and $Coll_{Smack}$, we may have found a relevant image suited as illustration. Usage Context information can also be useful for determining what the image, most likely, is not about.

4.2 Accumulating Usage Context descriptions

When an image is made publicly available, for instance on the Internet, complete Usage Context information may be stored in the header of the image file. This will make Usage Context information easily available for image retrieval purposes, and we avoid access to remote sites for locating and collecting context information.

As an image (or a copy of an image) may be copied several times and included in different collections or documents, the image metadata may include a number of Usage Context descriptions. As an example, consider figure 2, and assume that collections $Coll_{Boats}$ and $Coll_{Fishing}$ reside on site A. The author of the "Smack" document is searching for images to illustrate the text. She finds an interesting image, O, in collection $Coll_{Boats}$, and downloads a copy of O to site B. However, before O is downloaded, information describing all known context sources, i.e. collections $Coll_{Boats}$ and $Coll_{Fishing}$, are stored in the header of the image file.

When the image is included in the "Smack" document, this document represents a new Usage Context for the image, and Usage Context information, such as keywords from the abstract or from surrounding text, is added to the image metadata. If the image is made available for retrieval from site B, an image retrieval system will have access to information about all three Usage Contexts.

The example above illustrates that Usage Context information can be inherited from older copies of an image. The image copies are thus spread in a tree-like fashion, where images in the logical tree will store different sets of Usage Context information. Images in the leaf nodes will have the most advanced set of context information, and the different branches of the tree will accumulate different sets of context information.

The above approach is feasible as long as the header of the image file can be updated with new context information. For read-only image files, the context information must be stored in a file external to the image. A context repository, storing context information for a number of images, may be useful for this purpose.

Ideally, a complete set of Usage Context descriptions, including all Usage Context information for every copy of an image, should be available in the context repository. This implies that a context repository should contain all available context information for a set of images, and when a copy of an image is used in a new context, the repository is updated with new information.

5 Representing usage context information

We will in this section describe how Usage Context information can be organized and stored as metadata to images. To represent image context metadata we use RDF, and present a new context vocabulary, CTXT, that enables structuring of different types of image context information.

5.1 Using RDF

When an image is copied and moved to a new site, original metadata will follow the image copy, and new metadata, representing the new Usage Context, will be added. This means that context information must be represented so that it is extendible and understandable in different environments. RDF is a framework for resource description that supports such characteristics.

RDF is a W3C standard for encoding knowledge, and supports interoperability between applications or environments that exchange information. RDF also provides the ability to join data from a number of disparate vocabularies easily and consistently. Vocabularies can be created by different communities and groups as appropriate and mixed together as required, without needing any centralized agreement.

Existing vocabularies, such as DC and FOAF[10] (Friend of a Friend), do not offer the necessary elements to distinguish between context types as described in section III. We therefore define the CTXT vocabulary that includes useful constructs for describing image context information. The CTXT vocabulary can be used in RDF descriptions together with other RDF vocabularies defined elsewhere.

5.2 CTXT, a context vocabulary

The CTXT image context vocabulary will include constructs for describing both Usage and Creation Contexts. It represents a supplement to existing vocabularies, such as DC and FOAF, and is expected to be used together with other vocabularies in image metadata descriptions. CTXT uses XML syntax, adopts the conventions of RDF, and is inspired by both the Dublin Core (DC) and the FOAF vocabularies.

Figure 3 lists the vocabulary elements needed for description of Usage Context. Some of these elements, such as Description, Subject and Title, resemble elements found in the DC vocabulary. Other elements, such as Keywords and Terms, are new in CTXT.

[10] http://xmlns.com/foaf/0.1/

Element	What it means
Availability	Availability of the context. It can be public, private or restricted.
Description	A full-text description of the context.
Identifier	An ID value for the context. For public context this may be an URI.
Keywords	Descriptive terms specifically relevant for the context. Intended for human provided keywords.
Language	The language of the context information.
Subject	The topic of the context.
Terms	Automatically extracted (index) terms for the context.
Title	A name given to the context.
Type	Context type.

Figure 3. CTXT vocabulary elements

Class:ctxt:UsageContext
UsageContext - the environment where an image is used

Context types in this class:

Document	A file that contains information the user can view and/or hear. It includes text, image, drawing, sound, and other media types.
Collection	An aggregation of resources (for instance images)
SurroundingText	Text located near an image.

Figure 4. UsageContext class

We have defined a class `ctxt:UsageContext`, to describe different types of environments where an image can be used. Figure 4 shows the class and its context types. Currently, we have identified three Usage Context types that can be used with the `<Type>` element. By identifying context types, the CTXT vocabulary supports a diversified use of context information where Usage Context types such as `<Collection>` description and `<SurroundingText>` can be handled differently by the image retrieval system.

In figure 5 we show an example of context descriptions for one of the images depicted in figure 2. The image is included in two image collections and one document. For each collection, we have a `ctxt:UsageContext` description of type `<Collection>`. The image collection, named "Boats", is public, and we have both an URI to the collection and a detailed description of the collection. The other collection is private, and can only be reached by the collection owner(s). Therefore, there is no URI identifying this collection. However, the collection owner has agreed to contribute to the collective understanding of images, and has made the collection title and a short subject description of the collection available for the public.

The image is also included in a document titled "The Smack - its importance during 1900 - 2000". The context description includes both a description of the document, i.e. a `ctxt:UsageContext` of type `<Document>` description, and a representation of Surrounding Text, i.e. a `ctxt:UsageContext` of type `<SurroundingText>` description.

CTXT information can be combined with other RDF vocabularies to give a more complete RDF description of images. The context information presented in figure 5 can thus be part of a larger RDF description holding more information about the image.

```
<?xml version0"1.0" encoding="UTF-8"?>
<rdf:RDF
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:ctxt="http://www.caim.cs.uit.no/ctxt/0.1/"

    <ctxt:UsageContext>
        <ctxt:type>Collection</ctxt:type>
        <ctxt:title>Boats</ctxt:title>
        <ctxt:identifier>
            "http://www.caim.cs.uit.no/Boats"
        </ctxt:identifier>
        <ctxt:availability>public</ctxt:availability>
        <ctxt:description>....</ctxt:description>
        <ctxt:subject>.....</ctxt:subject>
    </ctxt:UsageContext>

    <ctxt:UsageContext>
        <ctxt:type>Collection</ctxt:type>
        <ctxt:title>Fishing in Northern Norway</ctxt:title>
        <ctxt:availability>private</ctxt:availability>
        <ctxt:subject>.....</ctxt:subject>
    </ctxt:UsageContext>

    <ctxt:UsageContext>
        <ctxt:type>Document</ctxt:type>
        <ctxt:title>
            The Smack - its importance during 1900 - 2000
        </ctxt:title>
        <ctxt:identifier>
            "http://www.caim.cs.uit.no/TheSmack.pdf"
        </ctxt:identifier>
        <ctxt:abstract>....</ctxt:abstract>
        <ctxt:subject>.....</ctxt:subject>
    </ctxt:UsageContext>

    <ctxt:UsageContext>
        <ctxt:type>SurroundingText</ctxt:type>
        <ctxt:identifier>
            "http://www.caim.cs.uit.no/TheSmack.pdf"
        </ctxt:identifier>
        <ctxt:keywords>....</ctxt:keywords>
    </ctxt:UsageContext>
</rdf:RDF>
```

Figure 5. Usage Context description

The RDF metadata can be saved as part of the comment block for JPEG, GIF, and PNG type images. A comment block in a JPEG file can contain arbitrary text, each block has a limited size, but there can be as many blocks as necessary. JPEG comment blocks are previously used in for instance [18] to store RDF data. RDF context information can also be stored in the Exif format. In that case, we are using the Comment element in Exif, that can include textual information of any kind. Independent of where the RDF data is stored, it should be expandable. As new image contexts emerge, the set of `ctxt:UsageContext` descriptions will be extended. We expect the CTXT vocabulary to expand. The CTXT vocabulary already has three Usage Context types available, and can be extended to support new types as needed. This will be represented as a new Type in the class `ctxt:UsageContext`. We also will expand CTXT with new classes, such as `ctxt:CreationContext`, representing information about the environment where the image was taken.

6 Conclusion

We have in this paper shown how Usage Context information, describing an image environment, can be used to indicate a semantic interpretation of the image. By combining information from multiple Usage Contexts, we enhance our knowledge of the image in that a number of semantic interpretations can be inferred, each representing different or complementary views of the image. Combined Usage Context information represents a new type of image metadata that can be used, together with other metadata types, in for instance image retrieval systems to improve retrieval strategies.

We present an RDF based context description vocabulary, called CTXT, that includes constructs for describing Usage Context information. CTXT can be combined with other RDF vocabularies to give a more complete description of images, including different types of image metadata. The CTXT vocabulary is extendable, since new context description constructs can be included as new image contexts are identified. Constructs for describing other image context classes, such as Creation Context, will be added, as well as more constructs for representing Usage Context.

Acknowledgements:
The authors appreciate support from the Norwegian Research Council, NFR, through project nr. 176858, Context-aware Image Management, CAIM, http://caim.uib.no

References

[1] Rong Zhao and William I. Grosky. Bridging the semantic gap in image retrieval. pages 14–36, 2002.

[2] M. L. Kherfi, D. Ziou, and A. Bernardi. Image retrieval from the world wide web: Issues, techniques, and systems. *ACM Comput. Surv.*, 36(1):35–67, 2004.

[3] G. Kowalski and M.T. Maybury. *Information Storage and Retrieval Systems - Theory and Implementation*. Kluwer Academic Publ, 2nd ed edition, 2000.

[4] Atsuo Yoshitaka and Tadao Ichikawa. A survey on content-based retrieval for multimedia databases. *IEEE Transactions on Knowledge and Data Engineering*, 11(1):81–93, 1999.

[5] A. Dey and G. Abowd. Towards a better understanding of context and context-awareness. In *Workshop on The What, Who, Where, When, and How of Context-Awareness, Conference on Human Factors in Computer Systems*, 2000.

[6] Adrian Graham, Hector Garcia-Molina, Andreas Paepcke, and Terry Winograd. Time as essence for photo browsing through personal digital libraries. In *JCDL '02: Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, pages 326–335, New York, NY, USA, 2002. ACM Press.

[7] Matthew Cooper, Jonathan Foote, Andreas Girgensohn, and Lynn Wilcox. Temporal event clustering for digital photo collections. *ACM Trans. Multimedia Comput. Commun. Appl.*, 1(3):269–288, 2005.

[8] Mor Naaman, Andreas Paepcke, and Hector Garcia-Molina. From where to what: Metadata sharing for digital photographs with geographic coordinates. In *10th International Conference on Cooperative Information Systems (COOPIS), LNCS 2888*. Springer Verlag, Berlin / Heidelberg, 2003.

[9] Y.-H. Lu and E. J. Delp, III. An overview of problems in image-based location awareness and navigation. In *Security, Steganography, and Watermarking of Multimedia Contents VI. Proceedings of the SPIE, Volume 5308*, pages 102–109, January 2004.

[10] Julien Pauty, Paul Couderc, and Michel Banatre. Using context to navigate through a photo collection. In *MobileHCI '05: Proceedings of the 7th international conference on Human computer interaction with mobile devices and services*, pages 145–152, New York, NY, USA, 2005. ACM Press.

[11] Mor Naaman, Susumu Harada, QianYing Wang, Hector Garcia-Molina, and Andreas Paepcke. Context data in geo-referenced digital photo collections. In *MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia*, pages 196–203, New York, NY, USA, 2004. ACM Press.

[12] F. Monaghan and D. OSullivan. Generating useful photo context metadata for the semantic web. In *Proceedings of the 7th International Conference on Mobile Data Management (MDM'06)*, 2006.

[13] M. Swain, C. Frankel, and V. Athitsos. Webseer: An image search engine for the world wide web, 1997.

[14] Thijs Westerveld. Image retrieval: Content versus context. In *Content-Based Multimedia Information Access, RIAO*, 2000.

[15] Linda Hill, Greg Janee, Ron Dolin, James Frew, and Mary Larsgaard. Collection metadata solutions for digital library applications. *Journal of the American Society for Information Science*, 50(13):1169–1181, November 1999.

[16] A. S. Chakravarthy and K. B. Haase. Netserf: Using semanticknowledge to find internet information archives. In *Proceeding of the ACM-SIGIR'95*, Seattle, Washington, United States, 1995.

[17] J. Xu and J. Callan. Effective retrieval with distributed collections. In *Proceedings of the 21st international ACM SIGIR conference*, Melbourne, Australia, 1998.

[18] Y. Lafon and B. Bos. Describing and retrieving photos using rdf and http. W3c note 19 april 2002, http://www.w3.org/tr/photo-rdf/, W3C, 2002.

[19] http://www.kanzaki.com/test/exif2rdf.

[20] Kerry Rodden and Kenneth Wood. How do people manage their digital photographs? In *CHI 2003 Conference on Human Factors in Computing Systems*, April 2003.