Context-Aware Media Retrieval

Ankur Mani[1], Hari Sundaram[1]

[1] Arts, Media and Engineering, Arizona State University, Tempe AZ 85282 (USA).
{Ankur.Mani, Hari.Sundaram}@asu.edu

Abstract. In this paper we propose a representation framework for dynamic multi-sensory knowledge and user context, and its application in media retrieval. We provide a definition of context, the relationship between context and knowledge and the importance of communication both as a means for the building of context as well as the end achieved by the context. We then propose a model of user context and demonstrate its application in a photo retrieval application. Our experiments demonstrate the advantages of the context-aware media retrieval over other media retrieval approaches especially relevance feedback.

Keywords: context modeling, media retrieval, knowledge representation

1 Introduction

In this paper we propose a representation framework for dynamic multi-sensory knowledge and user context, and its application in media retrieval. There are two fundamental problems in media retrieval: the vagueness of the query and the unavailability of the information about the unique way the user associates the different media elements. The context of the user provides solutions to both these problems at the same time and hence representation and estimation of user context is important.

We interpret the user's interaction with a media retrieval system as a special type of communication in which the user provides messages as query and the system provides messages as retrieved media. In any communication scenario, we define context as "*the finite and dynamic set of multi-sensory and inter-related conditions that influences the exchange of messages between two entities in communication*." This set forms a subset of knowledge that is "*a dynamic set of multi-sensory facts.*" Knowledge has three important properties; it is multi-sensory (represented through multiple senses), emergent (new facts are formed) and dynamic (old facts are revised). Context is the dynamic subset of knowledge that is in attention and influences the exchange of messages between the entities in communication. In a media retrieval scenario, the multi-sensory and interrelated information set in the user's short-term memory influences the query provided by the user [2] and at the same time is influenced by the user's activity and the media the user consumes. *This set of multi-sensory and interrelated information forms the user query context*. The organization of the rest of the paper is as such. We propose a user context model for media retrieval in section 2 based upon the above discussion. Section 3 discusses user experiments and finally we discuss conclusions in section 4.

2 Context for media retrieval

We now present our context model for media retrieval. The context model consists of two structures: a dynamic multi-sensory knowledge representation consisting of the concepts and the relationships between them, and a temporally evolving context representation in relation to this knowledge. The Knowledge can be subdivided into user knowledge (knowledge about the user), environment knowledge (knowledge about the environment, here limited to the common-sense knowledge from ConceptNet [3]) and application knowledge (media database structure consisting of media and related annotations).

2.1 Knowledge

Knowledge is represented as a graph. The nodes in the graph are the instances of concepts in one modality and the weighted edges (weights represent the similarity between the end nodes along the edge) are the relationships between those instances. The user knowledge model is initialized by an initial set of concepts obtained from the user as their profile and obtaining the neighborhood of this set in the environment knowledge. The user knowledge model grows as the user interacts with the media retrieval system that leads to information exchange between the user and the system. The environment knowledge and the application knowledge are represented similarly.

2.1 User query context

User query context is represented as the subset of the nodes and edges in the knowledge graph that are in attention. The attention on concepts and relationship types are represented as weights of the respective concepts and the biases on the types of relationships. The bias on the relationship type along with the weight of relationship determines the similarity between the neighboring concepts demonstrating the relationship. The user context evolves with the interaction and the weights of concepts and the biases on relationship types changes as discussed in [1]. Some important desirable properties of the context dynamics are the stability, controllability and suitable steady-state distribution of weights on concepts. It was proved in [1] that the dynamics has these properties.

3.1 Context-aware retrieval

We applied our context model to a photo retrieval system. The system is composed of four components as shown in Figure 1: a user context model, a media database, a user interface that allows users to enter text query or select relevant images and a search engine. Given the query as a set of selected images, the context-aware search is performed in the media knowledge space to find the most relevant photographs. The search process first obtains the current context from the context model and modifies it using the user information obtained from the query. The modified context is then used to obtain the candidate concepts in the media knowledge space. The images close to the candidate concepts in the media knowledge space form the retrieval results.

(a) (b)

Figure 1: (a) Block diagram of the photo retrieval application and (b) the user interface displays 9 images at a time and allows for query as text and images

3 Experiments

We now discuss the retrieval experiments. To compare the context-aware retrieval with baseline strategies, we performed experiments with three retrieval set-ups namely the random retrieval, relevance feedback based retrieval [4] and the context-aware retrieval. Six graduate students volunteered for the experiments and a database of approximately 4000 images (15% annotated) was made from their shared photo collection. Each user provided a set of at least ten concepts as the seed with which the user knowledge and context were initialized. Then each user searched for one query concept from among a set of choice concepts. Once the images were displayed, the user selected the relevant images that were used to retrieve new set of images without replacement. This process was repeated four times.

Table 1. Number of retrieved images and the mean relevance score for different queries and the % of relevant images in the database

Query	% database	Number of retrieved images; and Mean relevance score		
		Random	Relevance Feedback	Context
Home	10	6; 0.07	12; 0.29	14; 0.32
Birthday	10	4; 0.07	14; 0.24	12; 0.38
Park	20	8; 0.16	20; 0.44	23; 0.51
Office	5	1; 0.02	8; 0.13	10; 0.29

We analyze the experimental results as both cumulative precision of the overall retrieved set and the change in the relevance score with increasing interaction and the personal priorities of different users. We present the cumulative precision results as the number of relevant images that were retrieved in the complete experiment of five iterations and the mean relevance score of the retrieved images in the five iterations. The normalized relevance score for the retrieved set of N images is:

$$S = \frac{2}{N(N+1)} \sum_{j=1}^{N} (N + 1 - r_j) \; ; \text{ where } r_i = \begin{cases} i & i \text{ is relevant to query} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The cumulative precisions for three different search strategies are shown in **Table 1**. We observe that the context-aware retrieval gives better cumulative precision than the other retrieval strategies.

An important aspect of the context-based retrieval approach is that with increasing interaction more relevant images are retrieved as shown in Figure 2. The relevance score of the retrieved set is seen increasing with the increasing interaction in the context-based retrieval strategy. The relevance feedback based approach also shows and increasing trend but is not very consistent.

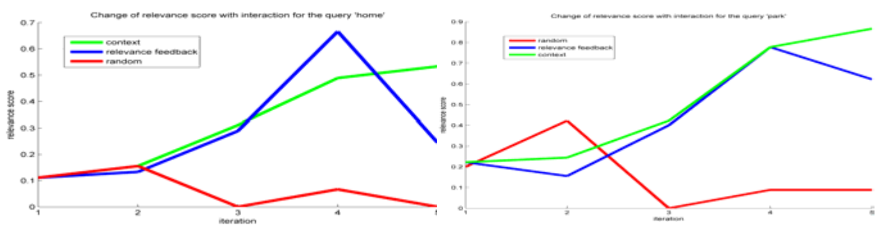

Figure 2: Plot of relevance score against user interaction for queries ‘home’ (top) and ‘park’ (bottom) queries.

The improved dynamic performance of the context-based approach against any other approach shows that with increasing interaction, the estimate of the user context becomes more accurate.

4 Conclusions

We presented a representation framework for dynamic multi-sensory knowledge and user context, and its application in media retrieval. Our pilot experiments demonstrated the advantages of context-aware media retrieval. In future along with more experiments, we also plan to expand the representation framework to support the representation of multi-scale and procedural knowledge.

References

1. Mani A., Sundaram H. Modeling User Context with Applications to Media Retrieval. to appear Multimedia Systems Journal, Springer Verlag. summer 2006.
2. Atkinson R. and Shiffrin R. (1968). Human memory: A proposed system and its control processes. The psychology of learning and motivation: Advances in research and theory(eds). New York, Academic Press.
3. ConceptNet http://web.media.mit.edu/~hugo/conceptnet.
4. Rui Y. and Huang T. (1999). A Novel Relevance Feedback Technique in Image Retrieval., Proc. ACM Multimedia 1999, Nov. 1999, Orlando, FL