

Multiple Viewpoints as an Approach to Digital Library Interfaces

James C. French A. C. Chapin Worthy N. Martin *
Department of Computer Science
University of Virginia
Charlottesville, VA
{french|acc2a|wnm}@cs.virginia.edu

1. INTRODUCTION

A viewpoint is a representational scheme on some collection of data objects, together with a mechanism for accessing this content. A multiple viewpoint system allows a searcher to pose queries in one viewpoint and then change to another viewpoint while retaining a sense of context, taking advantage of the improvement in retrieval results given by using different relevance judgements and different representations together in a single information retrieval system (as observed by Wilbur [6] and Rajashekar and Croft [4]).

The first advantage of a multiple viewpoint system is in providing a larger vocabulary of usable query terms. Formulating a query effectively depends on consistent use of vocabulary between the query and the document descriptions. A multiple viewpoint system can incorporate viewpoints that index the same collection in more than one vocabulary, accepting queries in either. A thesaurus which finds synonyms in other vocabularies is also an example of a viewpoint, and can further expand the usable query vocabulary.

The most important characteristic of a multiple viewpoint system is the transition mapping between each pair of viewpoints, which represents interrelations between data elements. For instance, if the same book is described in general science vocabulary in one viewpoint and in medical jargon in another, the transition mapping would map between these two representations allowing a searcher who had found the book in one viewpoint to see what was considered most similar to the same book in the other viewpoint.

Entirely different objects might also be associated by a transition mapping. If one viewpoint represented a document collection and another the relationships among a set

of authors, the transition mapping between these viewpoints might map book to author. A multiple viewpoint system that included viewpoints for many different media might allow a searcher to begin at Beethoven and, through various transitions, end up at the actor Malcolm MacDowall, the director Stanley Kubrick, or the writer Anthony Burgess (Figure 1).

2. CONTEXT

The concept of multiple viewpoints as we use it was described by Powell and French [3] as an approach to taking advantage of the findings of Wilbur [6], Rajashekar and Croft [4] and others regarding the potential for improving retrieval effectiveness.

Researchers such as Teraoka and Maruyama [5] have used the term “viewpoint” in representing a searcher’s interests and purpose, and “multiple viewpoints” in their work case consist of parameters to an information visualization system which indicate how to present information based on a particular (dynamically evolving) interest profile. While a viewpoint in our parlance might well be based on a particular user’s interests, we do not limit differences among viewpoints to different visualizations of the same information relationships; in our viewpoints, the relationships may differ as well.

Buckland’s Unfamiliar Metadata project [1] approaches the problem of mismatched vocabulary between searcher and index by suggesting query augmentation. An EVI (Entry Vocabulary Index) is a mapping from an ordinary language query to a list of possibilities drawn from the vocabulary used in the database. This is an example of using two viewpoints, the database itself indexed by its own vocabulary, and an “EVI viewpoint” which represents relationships between words in the two vocabularies, and whose output can be used as input to the other. A study demonstrating the potential of this approach can be found in French *et al.*[2].

3. DEFINING VIEWPOINTS

In describing a system of multiple viewpoints, we first identify the universe of data items with which we are concerned. This may include not only artifacts from one or more collections, such as the items in a database or books

*This work supported in part by NASA Grants NAG5-8585 and NAG5-9747 and NASA GSRP NGT5-50062.

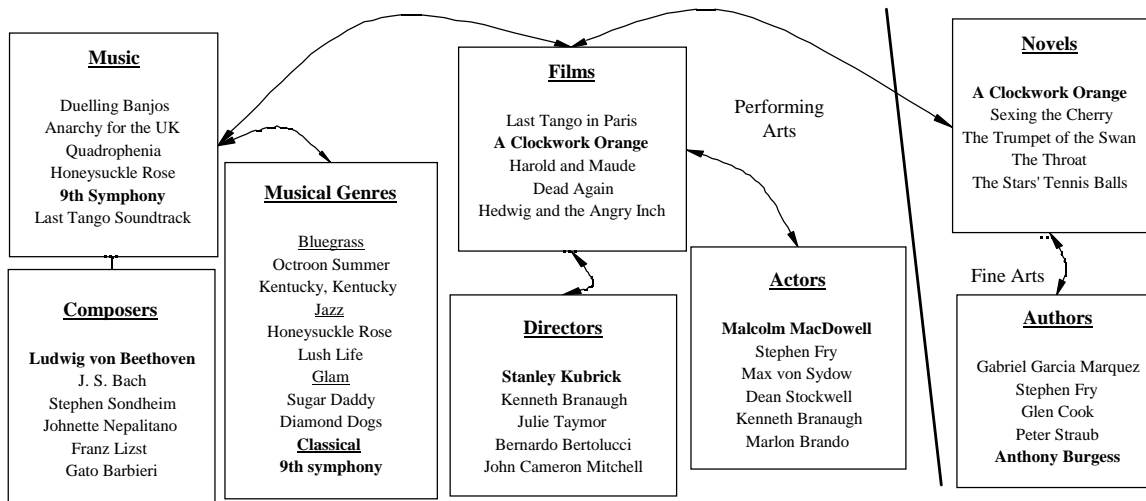


Figure 1: Viewpoints on various Media.

in a library, but also auxiliary data items such as keywords and saved queries.

Several data decompositions can be applied to the universe of data items, organizing it into categories; separating keywords from books is one such decomposition, but books might also be decomposed into fiction and nonfiction. Data decompositions are suitable for broad, rigid categorizations, but lack the fluidity of description available with viewpoints. A lens is a process which intersects some set defined in a data decomposition with another subset of the universe of data items; for instance, a lens for nonfiction books could be applied to the result of a library search so that only nonfiction books satisfying the search are shown.

Each viewpoint provides a representational scheme for some subset of the universe of data items, its viewpoint data set, and a mechanism for accessing this content. The set of possible inputs to the access mechanism are viewpoint queries. Each query has a corresponding viewpoint result, which is some subset of the viewpoint data set.

A text query to a multiple viewpoint system might have to undergo some translation and augmentation before it is usable as a viewpoint query for a particular viewpoint; the initial query transition performs this transformation. For each pair of viewpoints there is a transition mapping which, given a query in the first viewpoint, produces a query in the second viewpoint.

A result merge function takes all the viewpoint results that have been created from querying system viewpoints and organizes them into a system result suitable for displaying to the user. This may be done dynamically during the search process, showing the user how the result changes with different queries and different viewpoints, and may include or not include elements of any viewpoint result.

4. CONCLUSIONS

Multiple viewpoint systems take advantage of several

different sets of comparative judgments about some collection of data in retrieving information from a data collection.

Among the benefits we expect from multiple viewpoint systems are the expansion of available vocabulary for search queries, and the improvements in document retrieval observed in systems using more than one set of judgments about data. We also hope the concept of viewpoints will aid in the thoughtful design of systems to take advantage of these benefits.

5. REFERENCES

- [1] M. Buckland et al. Mapping Entry Vocabulary to Unfamiliar Metadata Vocabularies. In *D-Lib Magazine*. <http://www.dlib.org/dlib/january99/buckland/01buckland.html>, January 1999.
- [2] J. C. French, A. Powell, F. Gey, and N. Perelman. Exploiting A Controlled Vocabulary to Improve Collection Selection and Retrieval Effectiveness. In *Proc. Tenth International Conference on Information and Knowledge Management*, pages 199–206, 2001.
- [3] A. L. Powell and J. C. French. The Potential to Improve Retrieval Effectiveness with Multiple Viewpoints. Technical Report CS-98-15, Department of Computer Science, University of Virginia, 1998.
- [4] T. B. Rajashekar and W. B. Croft. Combining Automatic and Manual Index Representation in Probabilistic Retrieval. *Journal of the American Society for Information Science*, 46(4), 1995.
- [5] T. Teraoka and M. Maruyama. Adaptive Information Visualization Based on the User's Multiple Viewpoints – interactive 3D Visualization of the WWW. In *Proc. 1997 IEEE Symposium on Information Visualization*, 1997.
- [6] J. W. Wilbur. The Knowledge in Multiple Human Relevance Judgements. *ACM Transactions on Information Systems*, 16(2):101–126, 1998.