Collection Guiding: a New Framework for Handling Large Multimedia Collections

Stéphane Marchand-Maillet and Eric Bruno
Viper group - CVML - University of Geneva
http://viper.unige.ch/

Abstract. We propose a novel approach for the management of large multimedia collections. Our aim is to emphasize content management rather than document retrieval. In our framework, the collection itself is considered as a coherent entity bearing its own specific properties. We show that by characterizing optimal structures on the collection, we are able to create adaptive symbolic presentations of the collection under various viewpoints. We base on these visualizations our new exploration strategy where the user is literally guided within the collection content organized with respect to its inherent properties.

1 Introduction

We face a context where we need to automate the management of multimedia document collections as much as possible but where the presence of a human operator is made necessary to reach a sufficient level of efficiency and accuracy. The simplest example of a private user managing his/her own digital photo and video collection already calls for the use of a number of helper tools to efficiently keep track of all the content. Here [12], we look at multimedia information management in a "queryless" context. The user is faced with a (large) number of multimedia data. The tools we wish to describe should help the user in getting a comprehensive view of the content of these collections effectively. The base system is a simple view system that shows items in random order. Whereas this may look as a purely technical challenge, its development involves the comprehension of human perception of visual content and leads to problems that do not just find analytical solutions. From the context in which these tools should be developed, we look at what tools we have at hand for achieving our goal and assemble them in a unified context. This directs us towards the concept of "Collection Guiding" where the user embarks for a visit thru the multimedia space automatically created using state-of-the-art techniques for automated document analysis. Our approach re-locates the user at the center of the system and sets back the emphasis on Human-Computer Interaction.

2 Approach

Figure 1 details the structure of a content-based collection browsing system. It is interesting to see that a system designer may create several paths within the

graph created by the proposed components. The most intuitive system would go from the raw data to presenting the results to the user through the following path. First, pre-defined features are extracted from the documents. From there, via the definition of a similarity measure, inter-document distances are computed. Then, clustering or structuring is applied to grasp the underlying structure of the document collection. Then, in order to map the collection in a viewable space, a dimension reduction step projects each document as a point in 1-, 2- or 3D space.

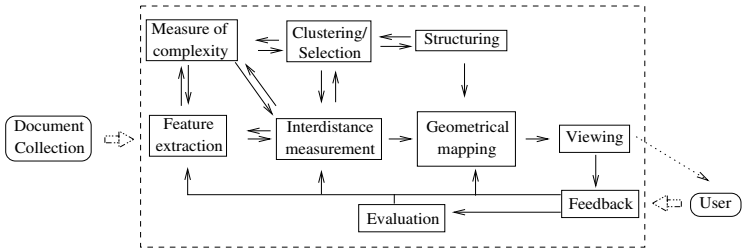

Fig. 1. Global structure

We also point out that most of the above system components are directly inherited from classical content-based data processing.

3 Constructing the collection visualization

3.1 Document processing

We consider items as represented by vectors in the feature space. Based on this principle, we may generalise our approach to any digital media, provided that suitable related features are defined. However, two aspects of this mathematical modeling should be inspected. First, we defined distances and similarity measures irrespectively of the feature space dimensionality. Yet, it is known that this dimensionality has an impact on the meaningfulness of the distances defined [1]. This is known as the *curse of dimensionality* and several results can be proven that show that there is a need for avoiding high-dimensional spaces, where possible. Further, typical visualization interfaces cannot handle more than 3 dimensions. Hence, there is a need for consistently representing items immersed in a high-dimensional space in lower dimensional spaces, while preserving neighboring properties.

3.2 Reduction of space dimension for visualization

Dimension reduction techniques aim, among other properties, at facilitating the inspection of abstract feature spaces [3, 11]. Techniques were initiated with Sammon's non-linear mapping [8] or Multidimensional Scaling (MDS [14]) and extended with curvilinear component analysis (CCA [5]), itself arising from the

study of self-organizing maps (SOM [9]). The common schema they follow is to summarize a set of points (x_i) in a D-dimensional space X with its inter-distance matrix $M_X = \{d_X(x_i, x_j)\}$ and to map in onto a set of points (y_i) such that the distance matrix $M_Y = \{d_Y(y_i, y_j)\}$ is as close as possible from M_X. A good review of techniques is given in [5].

3.3 Visualization strategies

The most basic 2D visualization interface is a simple grid of images organized into pages. Depending on whether the two dimensions are treated uniformly, this grid may for example be considered as a (vertical) list of (horizontal) series of items. In any case, here only the *order* of items presented is seen as important. In this case, the underlying structure is in fact essentially 1-dimensional. This visualization strategy is the most commonly used in search systems presenting ranked lists of items. Alternatives look at visualizing data into graph-based structures such as hierarchical tree. One typical area where such data visualization is used for Knowledge Base visualization [6]. Such strategies have also been exploited for displaying results of information search systems [2, 4, 7].

In this work, we are advocating for the use of 1D (or "near-1D") strategies to visualize (or visit) the content of a collection. The idea is that the collection should be organized along a route that shows an intuitive characteristic so as to be able to acquire easily the content of a collection. The "near-1D" principle is related to visiting a museum (hence the term "Collection Guide"). In a museum, although a user is following a linear path, (s)he encounters rooms within which paintings or objects are arranged. In other words, we look for a set of clusters arranged along a path. Alternatively, we look for an ordering that may not have a unit width. The main tool we wish to use here is the formulation of the Traveling Salesman Problem. This problem typically formalizes the exhaustive visit of a set of discrete locations along a route that shows some optimal properties.

3.4 Strategies for user interaction

Recent developments of multimedia systems have demonstrated the necessity and effectiveness of including user interaction within the browsing process [10]. Here, tools like relevance feedback [13] and online learning [15] remain valid. The use of incremental algorithms for solving optimality problems allow to have a good preconditionning of optimization problems and hence to reach a novel optimal solution rapidly upon modification of the constraints (*e.g.* via relevance feedback). Further, one should also note that the use of incremental algorithms allows to constrain computation time and to return "the best solution at time T", thus enhancing interactivity.

4 Conclusion

We have proposed and surveyed the development of the Collection Guide, both as a new principle and as a tool. The Collection Guide may apply to multimedia in

general. We show by a review of principles that the Collection Guide is a principle new to the Information Retrieval community. Its promises extend beyond what is presented here. In this sense, it is an important step forward towards the design of usable tools for multimedia content handling and description.

Acknowledgments

This work is supported by the Swiss National Science Foundation (SNF) via subsides number 21-66648-01 and 200020-105282/1 and the NCCR (IM)2.

References

1. Charu C. Aggarwal, Alexander Hinneburg and Daniel A. Keim. On the surprising behavior of distance metrics in high dimensional space. *Lecture Notes in Computer Science*, 1973, 2001.
2. M. Carey, D. Heesch and S. Rüger. Info navigator: A visualization tool for document searching and browsing. In *Proceedings of International Conference on Distributed Multimedia Systems*, September 2003.
3. M. A. Carreira-Perpinan. A review of dimension reduction techniques. Technical Report CS-96-09, Dept. of Computer Science, University of Sheffield, UK, 1996.
4. J.-Y. Chen, C. A. Bouman and J. C. Dalton. Hierarchical browsing and search of large image databases. *IEEE Transacrions on Pattern Analysis and Machine Intelligence*, 9(3), 2000.
5. P. Demartines and J. Hérault. Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets. *IEEE Transaction on Neural Networks*, 8(1):148–154, 1997.
6. Vladimir Geroimenko and Chaomei Chen, editors. *Visualizing the Semantic Web*. Springer Verlag, 2003.
7. M. A. Hearst and J. O. Pedersen. Reexamining the cluster hypothesis: Scatter/gather on retrieval results. In *ACM SIGIR'96*, Zurich, Switzerland, 1996.
8. J. W. Sammon Jr. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, 18:401–409, 1969.
9. T. Kohonen. *Self-Organizing Maps*, volume 30 of *Series in Information Sciences*. Springer, Berlin, 2001.
10. S. Kosinov and S. Marchand-Maillet. Overview of approaches to semantic augmentation of multimedia databases for efficient access and content retrieval. In *1st International Workshop on Adaptive Multimedia Retrieval (AMR2003)*, 2003.
11. J. X. Li. Visualization of high-dimensional data with relational perspective map. *Information Visualization*, 3:49–59, 2004.
12. Stphane Marchand-Maillet. Collection guiding: 1- principles. Technical Report 03.06, Viper Group - CVML - University of Geneva, 2003.
13. Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra. Relevance feedback: A power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Video Technology*, 8(5):644–655, 1998.
14. W.S. Torgeson. Multidimensional scaling of similarity. *Psychometrika*, 30:379–393, 1965.
15. N. Vasconcelos and A. Lippman. Learning from user feedback in image retrieval systems. In T.K. Leen S.A. Solla and K.-R. Müller, editors, *Advances in Neural Information Processing Systems 12*. MIT Press, 2000.