Interactively Browsing Large Image Collections

Ronald Richter, Mathias Eitz and Marc Alexa – TU Berlin

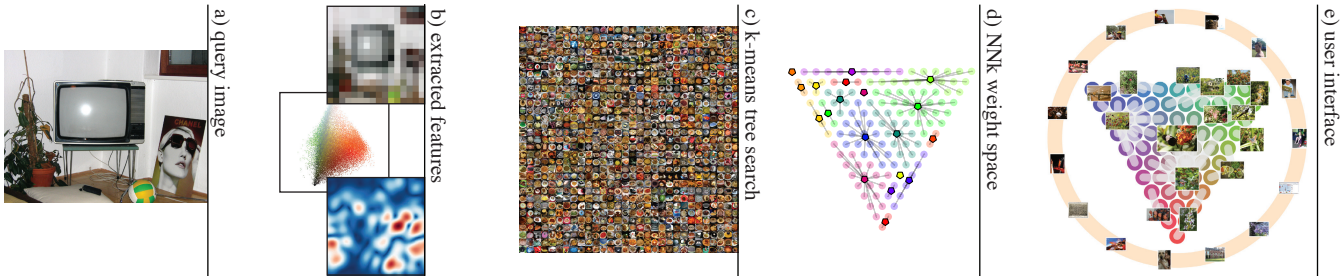

Figure 1: *Quick approximate nearest-neighbor search (c) is utilized to accelerate NNk search (d) yielding interactive image browsing (e).*

1 Introduction

Manually locating an image in a large collection has become infeasible with the recent rapid growth in size of such collections. Nowadays, even private collections easily contain tens of thousands of images; public collections have long passed the billion images mark. Current approaches for finding images in large collections, therefore, try to confine the set of images by returning only those images that correspond to certain properties defined by a query. Such properties can include: keywords, semantic information associated with the images, similarity to an example image, a rough sketch of the desired outlines, or any combination thereof.

However, common to all those methods is that they require the user to have an *a priori* notion of the image to be found. Very often, this is not the case and users only have a vague idea of the desired image or simply want to get a general idea of the variety of images contained in a collection. In this paper, we propose a novel method for interactively *browsing* large image collections, making the user an integral part of the interactive exploration by repeatedly exploiting the amazing ability of humans to quickly identify relevant images from a large set. The method requires only minimal input effort: users simply point to the image in the current display that seems most attractive to them (Fig. 1a)). The system then assembles a representative set of other images from the collection that are likely in its perceptual vicinity. The resulting browsing approach is – even for novice users – extremely simple to use and enables an interactive exploration of the collection as well as target-oriented selection towards a specific mental image model.

2 Interactive Browsing

The only data that is reliably available for *any* image collection is the data extractable from the images themselves. Consequently, we restrict ourselves to content-based methods and make use of existing image features such as texture information [Oliva and Torralba 2001], color distributions, and spatial color layouts (Fig. 1b)). When browsing images, a notion of image similarity is required. A simple definition is commonly given by defining a distance metric in image feature space. Images are then compared by computing distances between their corresponding feature vectors.

In the context of browsing, it is desirable to combine several different image features in order to cover multiple aspects of an image at once. However, the importance of individual features can differ according to the contents of a particular image as well as the intentions of the user. As we cannot predict the user's goal, a single, fixed linear combination of weights is inappropriate for combining features. Instead, NNk search has been proposed [Heesch and Rüger 2004], exploring the complete (quantized) weight-space (Fig. 1d)). The NNk of a query is defined as that set of images, that are nearest-neighbors to the query with respect to at least one combination of weights. NNk networks can be constructed offline – but the process is quadratic in the number of images in the collection and thus quickly becomes infeasible for large collections. In this work, we propose an approximation of NNk search that allows applying the underlying idea to theoretically arbitrarily large image collections without notable loss of quality. Instead of finding exact nearest-neighbors, we propose to restrict the search to promising candidates using an approximate k-means tree for sublinear searching [Fukunaga and Narendra 1975]. The resulting system reaches a precision of over 90% with respect to the full search but gives instant answers without the need for costly offline computation.

3 Results and Discussion

The resulting browsing interface – see Fig. 1e) – allows browsing large image collections interactively and efficiently. In the accompanying video, we demonstrate an *interactive* system using the proposed method on a collection of one million images. Users of the experimental system were able to quickly gain an overview of the collection as well as to quickly narrow down onto certain desired image categories. Using the proposed approximate NNk search, the system is theoretically scalable to web-scale collections, limited only by the amount of main memory available. Given the hierarchical nature of the underlying nearest-neighbor search datastructure, we believe that an out-of-core as well as a distributed version of the search (subtrees distributed over multiple machines) are natural candidates for future work. Additionally, we see potential for a *localized* version of the browsing interface, where users would fix parts of the image that are acceptable, e.g. using a multi-touch interface. Browsing would subsequently be restricted to those images that contain the fixed local features, exploring only the part of the collection that is relevant for the current search.

References

FUKUNAGA, K., AND NARENDRA, P. M. 1975. A branch and bound algorithms for computing k-nearest neighbors. *IEEE Trans. Computers 24*, 7, 750–753.

HEESCH, D., AND RÜGER, S. 2004. NNk networks for content-based image retrieval. *Advances in Information Retrieval*, 253–266.

OLIVA, A., AND TORRALBA, A. 2001. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *IJCV 42*, 3, 145–175.

SIGGRAPH 2010, Los Angeles, California, July 25 – 29, 2010.
ISBN 978-1-4503-0210-4/10/0007