Visual Berrypicking in Large Image Collections

Abstract
Exploring image collections using similarity-based two-dimensional maps is an ongoing research area that faces two main challenges: with increasing size of the collection and complexity of the similarity metric projection accuracy rapidly degrades and computational costs prevent online map generation. We propose a prototype that creates the impression of panning a large (global) map by aligning inexpensive small maps showing local neighborhoods. By directed hopping from one neighborhood to the next the user is able to explore the whole image collection. Additionally, the similarity metric can be adapted by weighting image features and thus users benefit from a more informed navigation.

Author Keywords
Interactive Exploration, Image Retrieval, Multi-Dimensional Scaling, Procrustes Analysis

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.; H.3.3 [Information storage and retrieval]: Information search and retrieval.

Introduction
When exploring large image collections, the user is interested in information about the collection as a whole:
what kind of images are available, how they relate to each other, and if there are clusters of similar images? In this scenario, similarity-based two-dimensional maps are a suitable means of visualization, because they quickly convey information about the general structure of a collection and help to identify groups of similar images. However, with growing size of the collection and increasing complexity of the underlying similarity metric, three major problems need to be faced:

• projections become increasingly inaccurate
• generating large maps is computationally impractical
• adapting feature weights (which enables personalized views) requires a recalculation of the distance matrix

In order to avoid these problems, we propose to visualize only the set of $k$-nearest neighbors for a given seed image in a small map. By choosing another image as a seed, the user is able to hop from one neighborhood map to another. Consecutive maps are likely to overlap to some extend and thus can be aligned to create a consistent transition that is, ideally, perceived as panning a large (global) map. This is visualized in Fig. 1. Because of this transition between two maps, users are able to transfer knowledge about the content and the relevance of individual images accumulated during the exploration process from one visualization to the next. This allows the user to navigate step-by-step through the whole collection, which we call visual berrypicking.

Related Work
In document-based information retrieval, berrypicking describes the user’s behavior during the search process [1]. Instead of a single query, the user performs a series of modified queries in order to find relevant information. In our scenario, choosing a neighboring image as a new seed corresponds to modifying a search query. An overview

![Figure 1: Images are projected to a two-dimensional map based on their pairwise similarity. The user hops from one neighborhood map to the next by selecting new seed images (with large border). Maps are aligned to create the impression of panning a large (global) map.](Image)
over different browsing models for image retrieval is given in [4]. Amongst others, the author discussed dimensionality reduction methods for the presentation of search results when addressing users with only poorly defined information needs. Rubner et al. [6] were among the first to propose multidimensional scaling (MDS [5]) for iterated image search in large collections – a technique that we also follow in this paper. The authors propose a local MDS on the nearest neighbors of a query image. However, in contrast to our approach, consecutive maps are not aligned to each other.

In previous work [8], we reviewed and compared different dimensionality reduction algorithms for the visualization of large music collections. In a user study, MDS was favored by most participants as best layout algorithm when the collection undergoes changes due to newly added items. Different from the approach proposed here, we pursued a global map approach for MDS. We also suggested to use Procrustes analysis [3] to better align newly generated maps with their respective predecessors, which we adopt here as it reduces confusion for users.

**Map Generation and Transitions**

Visualization of image similarities on a two-dimensional map requires dimensionality reduction of the typically much higher input feature space. We used MDS for that purpose. Despite providing the advantage of being distance preserving, MDS as any projection into lower dimensional spaces will cause projection errors that increase with the number of dimensions to reduce and the size of the collection. This also applies to one-dimensional lists, i.e., the visualized order of objects may not appropriately respect the order defined by the used metric. By limiting the number of images used to compute the projection, we reduce the impact of this error and thus visualizations become more reasonable.

Transitions between consecutive maps are animated with the aim of giving the user the impression of panning a large map representing the collection as a whole. In order to make these transitions as consistent as possible, we align consecutive maps on their common neighbors. We use Procrustes analysis to reduce the sum of the squared differences between the two sets of images that remain...
visible by translation, scaling, rotation and reflection. As a result, the user benefits from continuity that allows to transfer knowledge from one map to the next and more stable navigation directions and, thus, is less likely to get lost.

**Interface Prototype**
We have implemented a web-based interface prototype that supports a traditional list-based as well as the proposed map-based approach for image collection exploration, see Fig. 2 (left and right). The prototype renders thumbnails of the \( k \)-nearest neighbors of a given seed image. By clicking on a thumbnail, the user can select a new seed and thereby retrieves another subset of images most similar to the one selected. This enables the user to navigate to the desired image or explore the collection as a whole. Since very similar images are likely to overlap in the MDS based visualization (see Fig. 2 right), we implemented a grid layout that preserves similarities and, in addition, avoids overlaps, see Fig. 3.

In our prototype, we used the Caltech101 data set [2] comprising 9144 images representing 101 categories. The employed similarity metric is based on a weighted sum over four individual visual image features: we extract the MPEG-7 descriptors Color Layout, Scalable Color and Edge Histogram [7] as well as a YCbCr Color Histogram. During exploration, the user is able to adapt feature weights to influence navigation directions.

**Conclusions**
We implemented a prototype to demonstrate our proposed interaction scheme for exploring large image collections. Instead of computing a large costly map for the whole collection we propose to generate small maps using MDS showing local neighborhoods. By selecting a new seed image, the user is able to hop from one neighborhood to another. Consecutive maps are aligned using Procrustes analysis in order to create the impression of panning a large (global) map. A preliminary user study showed that users benefit from additional information provided by aligned maps. In a next step, we plan to conduct a user study that compares both list- and map-based visualizations in more detail. Also, we will investigate how to increase the overlap between consecutive maps such that the alignment can be improved.

**References**