Exploring Large Movie Collections: Comparing Visual Berrypicking and Traditional Browsing

Thomas Low[1], Christian Hentschel[2], Sebastian Stober[3], Harald Sack[4], and Andreas Nürnberger[1]

[1] Otto von Guericke University Magdeburg, Germany
{thomas.low,andreas.nuernberger}@ovgu.de
[2] Hasso Plattner Institute for Software Systems Engineering, Germany
christian.hentschel@hpi.de
[3] University of Potsdam, Germany
sebastian.stober@uni-potsdam.de
[4] FIZ Karlsruhe, Germany
harald.sack@fiz-karlsruhe.de

Abstract. We compare Visual Berrypicking, an interactive approach allowing users to explore large and highly faceted information spaces using similarity-based two-dimensional maps, with traditional browsing techniques. For large datasets, current projection methods used to generate maplike overviews suffer from increased computational costs and a loss of accuracy resulting in inconsistent visualizations. We propose to interactively align inexpensive small maps, showing local neighborhoods only, which ideally creates the impression of panning a large map. For evaluation, we designed a web-based prototype for movie exploration and compared it to the web interface of *The Movie Database* (TMDb) in an online user study. Results suggest that users are able to effectively explore large movie collections by hopping from one neighborhood to the next. Additionally, due to the projection of movie similarities, interesting links between movies can be found more easily, and thus, compared to browsing serendipitous discoveries are more likely.

Keywords: exploratory interfaces, media retrieval, multidimensional scaling, user study

1 Introduction

Exploring movie collections containing thousands of items described by textual metadata such as title, genre, actors and plot summary is difficult for users, as it takes some effort to judge the relevance of all retrieved objects. In this paper, we present and evaluate a map-based interface in order to support especially exploratory processes when working with highly faceted information spaces. Current projection methods suffer from increased computational costs and a loss of accuracy resulting in inconsistent visualizations. However, previous work has shown that similarity-based two-dimensional maps can be very useful for exploratory information retrieval [12]. They quickly convey information about the general structure and coverage of a collection and help to identify groups of similar objects. Grouping by similarity helps to assess the relevance of entire clusters once the relevance of a few representative objects is known.

In order to evaluate the effectiveness of our proposed map-based exploration method, we compared it with standard browsing techniques when exploring a large collection of movies. Our exploration model enables a user to quickly identify movies that share common features such as common actors or similar plots. Navigation through the collection is possible by panning towards a specific feature subset.

In [11] the conceptual idea of Visual Berrypicking was demonstrated by extracting visual features from a large image collection. In this paper, we transfer and evaluate this approach for a more complex movie dataset by combining its diverse metadata, e.g., a movie's rating, plot description and more. We follow the approach of [11] and visualize only the set of k-nearest neighbors for a given reference or *seed* object. Selecting a new seed from the currently visualized subset leads to the retrieval of a new set of k-nearest neighbors and the computation of a new map. The old and the new set of visualized neighbors most likely overlap to some extent. This topological similarity is exploited to align the two consecutive maps and create a meaningful link between them. Additionally, animated transitions aim to help keeping track of positional changes for objects contained in both maps. Ideally, with largely consistent transitions, this creates the impression of panning a large (global) map of the collection. Most importantly, establishing links between consecutive maps allows users to transfer knowledge about the content and relevance of individual objects accumulated during the search process from one visualization to the next.

For evaluation, we conducted an online user study with more than 100 participants. Participants were asked to explore the collection of *The Movie Database*[5] (TMDb) and find interesting movies they were eager to watch using both the native website as well as our proposed map-based movie explorer. Details about our study as well as its results are discussed in Section 4. First, we briefly review related work in Section 2 and present details of the Visual Berrypicking method in Section 3.

2 Related Work

Media retrieval systems are based on either or both extracted audiovisual features and other metadata. Approaches that focus on extracted audiovisual features usually try to tag movies or individual scenes and provide faceted search interfaces. As an example, the authors in [6] present a semantic video search engine that enables shot-accurate exploration of a cultural heritage video archive. While the interface allows for filtering retrieval results using content-based facets, a query refinement needs to be performed by altering the initial query string.

A dense visualization of a large collection of videos is presented in [1]. Browsing is supported by tree-based navigation, each level representing a two-dimen-

[5] The Movie Database Website: `https://www.themoviedb.org/`

sional projection of similar sequences. While this extremely compact layout maximizes use of available display space, identifying clusters of similar elements becomes difficult and impossible when groups are located in different hierarchy levels. Furthermore, exploration is limited by navigating up and down branches of the collection tree. Different from the aforementioned approaches, we do not further analyze movies. Our approach solely relies on metadata provided by TMDb and therefore can be easily applied to other media collections, e.g., images, music or simple text documents.

Rubner et al. [13] were among the first to propose multidimensional scaling (MDS) for iterated search in large image collections – a technique that we also follow in this paper. The authors propose a local MDS on the nearest neighbors of a query image. In contrast to our approach, consecutive maps are not aligned to each other. In [12] image arrangements are constructed using MDS and layouted continuously as well as in a structure preserving grid layout (in order to reduce overlap). The authors conclude that arranging a set of thumbnails according to their similarity is useful and helps to divide the set into simple clusters. Several other map-based approaches for information retrieval have been described, e.g. for large document collections using self-organizing maps [8], graph-based approches [5] or adaptive multi-view systems [10].

Recent work [14] reviewed and compared different dimensionality reduction algorithms for the visualization of large music collections. Based on a user study, MDS was favored as best layout algorithm when the collection undergoes changes due to newly added items. In contrast to this paper, [14] pursued a global map approach for MDS. It was suggested to use Procrustes analysis [4] to better align newly generated maps with their respective predecessors, which we will also adopt here as it leads to a reduced confusion of users when navigating with aligned maps. Similar to our approach, Kleiman et al. [7] arrange images on a fully populated similarity based grid layout that can be interactively panned and zoomed. Due to the dense layout, similarities and clusters cannot be inferred from the layout alone. In case of images, humans are able to quickly assess similarities based on the visual information. When using other media, map-based approaches need to illustrate similarities and clusters in order to support the user.

Evaluating map-based retrieval and exploration systems is very challenging due to the complexity of the exploration process. Often, systems are evaluated in controlled lab experiments in order to get a better understanding of how users interact with them and whether there are able to utilize the two-dimensional arrangement. In [12] random and similarity based organization of images in browsing scenarios are compared. Strong et al. [15] evaluate the user's effectiveness in finding images with specific properties. These evaluation scenarios do not conform to a typical exploration task, where a user's information need might change during the exploration process. In this paper we aim to evaluate a user's exploration effectiveness while looking for movies that are considered to be worth watching tonight. To our knowledge, there is no online study comparing a map-based exploration approach to an established, traditional browsing-based system that evaluates the user's effectiveness in an open exploration scenario.

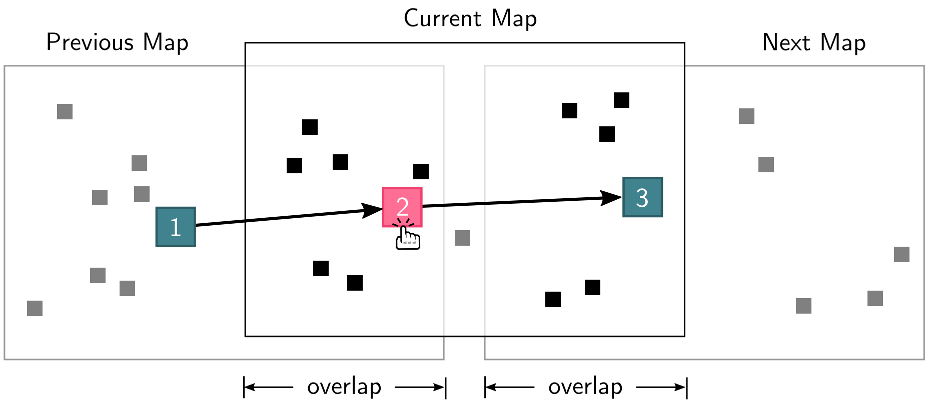

Fig. 1. Similarity-based projection of nearest neighbors (squares) using three seed items (colored squares 1,2,3). Common neighbors (black squares) overlap between consecutive maps and are used for alignment when navigating from one item to the next.

3 Visual Berrypicking

In document-based information retrieval, *berrypicking* describes the user's behavior during the search process [2]. Instead of a single query, the user performs a series of evolving queries in order to find relevant information. While inspecting individual documents, the user gets a better understanding of his or her own information need, which is then used to modify the query. At any time during this process, useful information can be identified, which will all contribute to satisfying the user's information need. When applying the idea of berrypicking to map-based visualizations, choosing a neighboring information or movie object as a new seed corresponds to modifying a search query, which we call Visual Berrypicking. We hypothesize that being able to iteratively inspect parts of the information space by Visual Berrypicking stimulates exploration, which helps to learn about information objects and their relations, and thus, enhances overall user experience.

Visualization of similarities on a two-dimensional map requires dimensionality reduction of the typically much higher input feature space. Multidimensional scaling (MDS, [9]) is a popular distance-preserving technique that can be used for this purpose. Naturally, 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. As a result, neighboring objects may turn out to be not that similar after all (degrading *trustworthiness*[2]) and similar objects may not be visualized as neighbors but far apart from each other (degrading *continuity*[6]). Such problems may disturb the process of berrypicking, since users might get

[6] The measures of trustworthiness and continuity are introduced in [16] together with a discussion of common problems that arise from using map visualizations for information retrieval.

confused about information objects that are visualized far apart though being perceived as similar. By limiting the number of items used to compute the projection, we try to reduce the impact of these problems. Large-scale collections with millions of items cannot be reasonably handled in their entirety anyway and users will therefore always focus on a small subset. For each view of the dataset, our prototype presents a map of the k most similar items only, given a user defined seed item. By clicking on any of the presented items, a new map is created using the selected item as a new seed.

Transitions between consecutive maps are animated with the aim of giving the user the impression of panning on a large map representing the collection as a whole. In order to make these transitions as consistent as possible, we use the overlap between any two consecutive maps to align them based on their common neighbors, see Figure 1. We use Procrustes analysis [4] to reduce the sum of the squared differences between the two sets of items that remain visible by translation, scaling, rotation and reflection. The alignment error corresponds to the difference between the two-dimensional positions of common nearest neighbors in consecutive maps after alignment by Procrustes analysis. Because of small relative position changes of common neighbors and the alignment of subsequent maps by Procrustes analysis this transition is ideally perceived as panning a structurally stable map. As a result, the user benefits from continuity that allows to transfer knowledge from one map to the next and more stable navigation directions (items with certain properties can be found in the same corner during multiple interactions). Thus, the user is less likely to get lost, which supports the process of exploration.

4 Evaluation

For evaluation we implemented a web-based prototype for movie exploration called NEMP (Neighborhood Exploration using MDS and Procrustes analysis) and compared it to the web interface of *The Movie Database* (TMDb), see Figure 2 and 3. We used the 10,000 most popular movies from TMDb, including a movie's title, cover image, rating, genres, actors, directors and plot description. Because the proposed technique only considers neighborhoods of a constant size for map generation, it does not depend on the collection size and thus is easily scalable – given that the k-nearest neighbors can be retrieved efficiently.

For computing movie similarities, we linearly combined five individual measures: linear difference in release date, jaccard similarity of genres, directors and actors, and plot similarity using tf-idf. Optimal weights were determined based on preliminary test trials. The number of k-nearest neighbors was fixed to 30. All pairwise similarities were calculated before the experiment, such that the k-nearest neighbors can be retrieved in $\mathcal{O}(n)$ time. However, for larger datasets there are more efficient ways to retrieve the k-nearest neighbors, e.g. approximately via locality-sensitive hashing [3].

Our map-based layout of movie items is constructed by applying classic MDS to the given seed item and all its k-nearest neighbors. Consecutive maps are

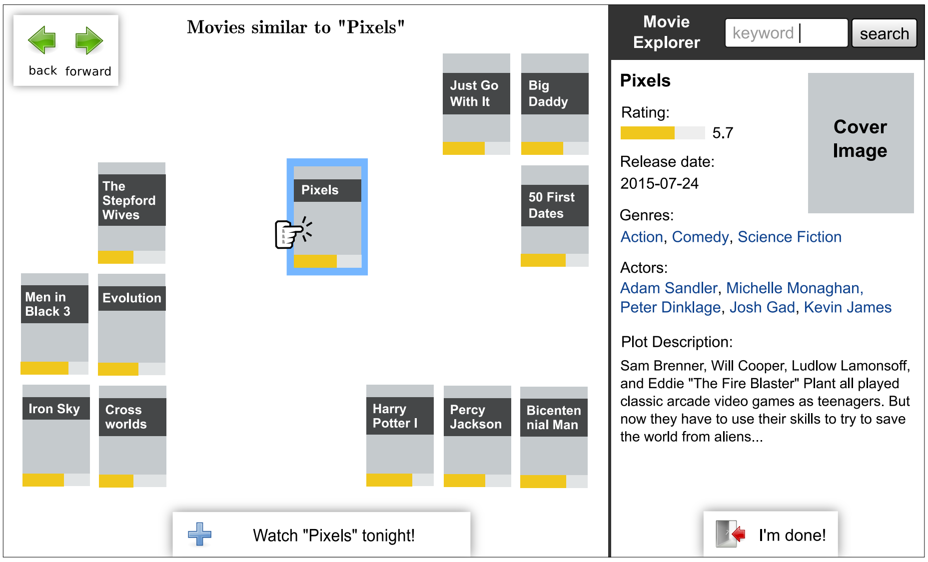

Fig. 2. NEMP User interface and study overlay (buttons with shadow): a projection of movies visualizes similarities and local clusters with respect to, e.g., genre (left bottom cluster), actors (top right cluster) and director (bottom right cluster); back and forward buttons (top left) as well as buttons to add a movie or quit are used as an overlay in both study interfaces. Cover images (grey areas) are omitted for legal reasons.

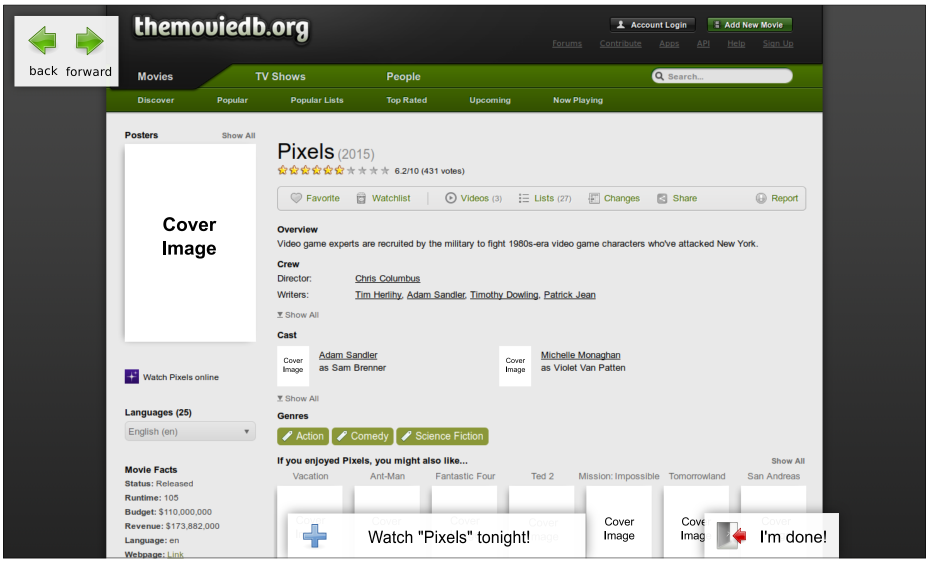

Fig. 3. TMDb user interface and study overlay: movies can be explored by inspecting lists of popular movies or following recommendations on movie pages. Screenshot adapted to fit page, cover images omitted. Original version at `https://web.archive.org/web/20151003235059/https://www.themoviedb.org/movie/257344-pixels`.

aligned using Procrustes analysis. In order to make use of all available display space, item coordinates are transformed linearly to better fit the screen's aspect ratio. Also, overlaps of similar movies are prevented by dividing the viewport into grid cells of fixed size and assigning movies to their closest free cell. Due to this transformation, the arrangement in the user interface slightly differs from the coordinates provided by MDS. We believe that the gain in usability outweighs the loss in precision of relative distances. Most importantly, clusters of similar movies are usually preserved.

The user interface is composed of a simple search bar, a sidebar showing detail information, and a large viewport presenting movie covers arranged in our two-dimensional layout, see Figure 2. By clicking on a movie cover, a new similarity search is started and its k-nearest neighbors are shown. Changes in the position of movies are animated. Hovering over a movie cover allows to inspect its information inside the details panel. Users are able to click on genres, actors and directors, which will result in a corresponding search query. Also, genres, actors, directors and plot terms that are common to both – the currently selected and hovered movie – are highlighted.

In contrast, the TMDb web interface follows a traditional browsing approach, see Figure 3. Users may browse through several lists of popular movies, top rated movies or movies starring a specific actor. Each movie can be inspected in a separate details page, which also provides a list of related movies as recommendations. In comparison to NEMP, more information are presented for each movie, including, e.g., its runtime, budget, additional pictures and reviews. Both interfaces allow to search for movies using keywords.

4.1 Study Design

We conducted a web-based online study comparing both NEMP and TMDb in an interactive exploration session. The task was to explore TMDb's movie collection in order to find movies a participant would consider worth watching tonight. Therefore, we asked participants to add movies they thought to be worth watching tonight to their personal watch lists. Our hypothesis is that users are better supported in exploring large movie collections by the proposed interface, resulting in more movies being added to watch lists.

In order to effectively test our hypotheses, we tried to avoid that users will add movies they have not found during the process of exploration. Therefore, we asked them to write down movie titles they could easily remember and would consider worth watching tonight before using any of the two websites. We assumed that participants would prefer to explore movies given that we asked them not to add movies they were aware of beforehand. The study interface itself was designed as consecutive pages and forms asking participants for 20–30 minutes of their time, their age, gender, profession and a self-assement of their own "movie knowledge". After a number of instructions, e.g., to use the fullscreen mode of their browser window, they were presented both user interfaces in random order (later denoted as *first* and *second* interface). During both sessions we asked participants to add relevant movies to their personal watch

helpful	I have the impression I was able to find interesting movies.
easy	I thought the interface was easy to used.
complex	I found the interface unnecessarily complex.
intuitive	I found navigation using this interface intuitive.
inconsistent	I thought there was too much inconsistency when new movies were shown.
interesting	I was able to find interesting links between different movies.
random	The movies presented seemed random.

Table 1. Evaluation statements that were rated according to a 5-point Likert scale by all participants immediately after using an interface.

lists. Therefore, we overlayed both interfaces by additional buttons that allowed to add and remove movies, as well as to proceed with the study once finished, see Figure 2 and 3. Immediately after each interface session, participants were asked to rate statements listed in Table 1 using a 5-point Likert scale.

4.2 Study Results

A total of 110 participants (48 female, 62 male, 31.8 years old on average) completed our study. Since participants performed our study online without any human supervision, results are unsurprisingly diverse. Although we implemented an automatic check for browser size, which shows a large green tick or red cross indicating whether the browser size is appropriate, a total of 67 participants chose to proceed with the study even though their window was rated to be too small. However, a large window size is important in order to provide enough display space to be able to identify clusters using the NEMP interface. Therefore, it would be interesting to compare our results with a controlled lab experiment.

In 9 cases, participants spent less than one minute in one or both of the two interfaces. Also, results suggest a strong bias towards the interface presented first, e.g., the average time spent on the first interface is 7:09 (min:sec) compared to 4:56 for the second interface. Due to the design of our online study, it was not possible to even out the number of trials starting with a particular interface. In total, 47 participants started with NEMP, 63 with TMDb. In order to do meaningful statistical comparisons, we randomly removed 16 results for participants that started with TMDb.

Figure 4 shows aggregated ratings for all statements of Table 1. Although results do not show a clear winner, some important aspects clearly stand out. The most prominent difference was expressed for the statement that interesting links can be found more easily using NEMP, which is the fundamental goal of our exploration interface and supports our initial hypothesis that users will be supported in learning about information objects and their relations.

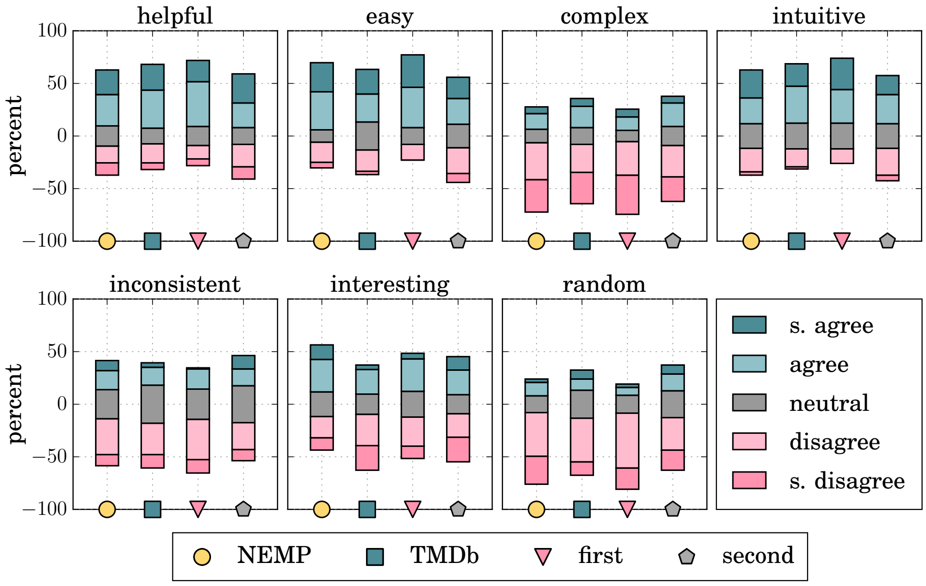

Fig. 4. Evaluation results presented in diverging stacked bar charts showing percentage of participants rating both interfaces according to all seven statements, see Table 1. Results for the first and second interface are shown for comparison.

Presumably due to the inherent differences in both interfaces, e.g., TMDb contains a lot more information, images, etc. about each movie, NEMP is rated slightly less helpful in finding interesting movies, but also more easier to use and less complex. Also, we found strong correlations between the participants' browser size and perceived complexity. TMDb was rated more complex with increasing browser size (Pearson correlation coefficient $pcc = 0.25$), while NEMP was rated less complex ($pcc = -0.12$), see Figure 5. Similarly, we found a positive correlation between NEMP's helpfulness and browser size ($pcc = 0.16$), which might be due to NEMP's two-dimensional presentation, which benefits from large screens. Finally, there was a strong correlation between the time participants spent using NEMP and their perceived randomness of presented movies ($pcc = -0.21$), suggesting that participants needed to get familiar with our approach.

On average 5.08 (standard deviation $std = 4.45$) movies were added to watch lists during exploration sessions using NEMP, and 5.77 ($std = 5.17$) were added during sessions using TMDb. Therefore, our hypothesis that more movies will be added while using the proposed exploration interface can not be validated. However, since NEMP is considered to be an unconventional user interface, it could still be clearly demonstrated that participants were able to effectively use it for exploration of a large movie dataset even though they used it for the first time.

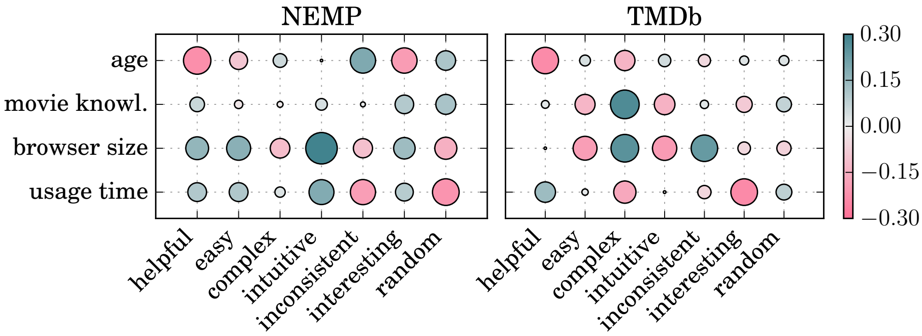

Fig. 5. Pearson correlation coefficients for study parameters and statement ratings of both interfaces, see Table 1. Circle radius and color intensity corresponds to correlation strength, hue to orientation.

5 Conclusions

We have compared Visual Berrypicking to traditional browsing techniques using a large-scale movie collection. Local maps are aligned during navigation from one neighborhood to the next, which ideally creates the impression of panning a large global map. At the same time, small maps can be computed much quicker and have higher visual accuracy. We have presented a web-based prototype for movie exploration using multidimensional scaling for map generation and Procrustes analysis for alignment. For evaluation, we have compared our prototype with the website of *The Movie Database* in an interactive online study. Results indicate that users can effectively explore large movie collections, even though our approach was unfamiliar to them. Additionally, they find more interesting links between different movies.

Although our study was conducted using the example of movie exploration, the proposed approach is not restricted to a specific application. Given a meaningful similarity measure and thumbnails (e.g. images or music cover art) our approach is generally applicable to any kind of media. Future work will focus on training feature weights during the user interaction process in order to further adapt maps to the user's retrieval focus. As local neighborhood maps are easily computed on the fly, adaptations of the underlying similarity space can be visualized immediately.

References

1. Barthel, K.U., Hezel, N., Mackowiak, R.: Graph-based browsing for large video collections. In: Proc. of the 21st Intl. Conf. on Multimedia Modeling. pp. 237–242. MMM '15, Springer (2015)
2. Bates, M.J.: The design of browsing and berrypicking techniques for the online search interface. Online Information Review 13(5), 407–424 (1989)

3. Gionis, A., Indyk, P., Motwani, R., et al.: Similarity search in high dimensions via hashing. In: VLDB. vol. 99, pp. 518–529 (1999)
4. Gower, J.: Generalized procrustes analysis. Psychometrika 40(1), 33–51 (1975)
5. Gretarsson, B., Odonovan, J., Bostandjiev, S., Höllerer, T., Asuncion, A., Newman, D., Smyth, P.: Topicnets: Visual analysis of large text corpora with topic modeling. ACM Transactions on Intelligent Systems and Technology (TIST) 3(2), 23 (2012)
6. Hentschel, C., Hercher, J., Knuth, M., Osterhoff, J., Quehl, B., Sack, H., Steinmetz, N., Waitelonis, J., Yang, H.: Open up cultural heritage in video archives with mediaglobe. In: Proc. of the 12th Intl. Conf. on Innovative Internet Community Services. I2CS (2012)
7. Kleiman, Y., Lanir, J., Danon, D., Felberbaum, Y., Cohen-Or, D.: Dynamicmaps: Similarity-based browsing through a massive set of images. In: Proc. of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 995–1004. CHI '15, ACM (2015)
8. Kohonen, T., Kaski, S., Lagus, K., Salojarvi, J., Honkela, J., Paatero, V., Saarela, A.: Self organization of a massive document collection. IEEE transactions on neural networks 11(3), 574–585 (2000)
9. Kruskal, J.: Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29, 1–27 (1964)
10. Lee, H., Kihm, J., Choo, J., Stasko, J., Park, H.: iVisClustering: An interactive visual document clustering via topic modeling. In: Computer Graphics Forum. vol. 31, pp. 1155–1164. Wiley Online Library (2012)
11. Low, T., Hentschel, C., Stober, S., Sack, H., Nürnberger, A.: Visual berrypicking in large image collections. In: Proc. of the 8th Nordic Conf. on HCI. pp. 1043–1046. NordiCHI '14, ACM (2014)
12. Rodden, K., Basalaj, W., Sinclair, D., Wood, K.: Does organisation by similarity assist image browsing? In: Proc. of the SIGCHI Conf. on Human Factors in Computing Systems. pp. 190–197. CHI '01, ACM (2001)
13. Rubner, Y., Guibas, L., Tomasi, C.: The earth mover's distance, multi-dimensional scaling, and color-based image retrieval. Proceedings of the ARPA Image Understanding Workshop pp. 661–668 (1997)
14. Stober, S., Low, T., Gossen, T., Nürnberger, A.: Incremental visualization of growing music collections. In: 14th Intl. Conf. on Music Information Retrieval. pp. 433–438. ISMIR '13 (2013)
15. Strong, G., Hoeber, O., Gong, M.: Visual image browsing and exploration (vibe): User evaluations of image search tasks. In: Intl. Conf. on Active Media Technology. pp. 424–435. Springer (2010)
16. Venna, J., Kaski, S.: Local multidimensional scaling. Neural Networks 19, 889–899 (2006)