

Exploring Online Video Databases by Visual Navigation

Wolfgang Hürst
Utrecht University
Utrecht, Netherlands
huerst@uu.nl

Bruno dos Santos Carvalho
Utrecht University
Utrecht, Netherlands
b.dossantoscarvalho@uu.nl

ABSTRACT

We present an interface design for interactive exploration of large movie databases based on a concept we entitle *visual navigation*. Our approach aims at combining the major advantages of existing systems, which are commonly either simple but limited in functionality or powerful but complex and less engaging. To verify the potential of our idea, we performed a pilot study, which indicates the validity of our approach, highlights advantages, and pinpoints areas for improvement and future work.

Author Keywords

Video database browsing, video archive access, movie database interfaces, interaction design, user experience.

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation (e.g. HCI)]: User Interfaces – Graphical user interfaces (GUI), input devices and strategies, interaction styles, screen design

INTRODUCTION & VISUAL NAVIGATION CONCEPT

Online databases such as *Netflix* (<http://www.netflix.com>) provide users with instant access to a tremendous number of movies and TV shows. Yet, the amount of available data results in challenges for the interface design. Common interfaces rely on a rather simplistic representation – likely to not overwhelm users and keep the interaction simple. They generally offer a somehow categorized overview of movie posters, which is complemented by a detailed view of a movie once it is selected from the overview; cf. Fig. 1. We refer to this type of video database browsing as *visual browsing*, because it enables users to browse and access movies based on visual information; a movie’s poster.

This works well in situations where the recommendations provided by the interface fit a user’s current needs. Yet, these needs can also change frequently based on status, mood, or context. Two common issues in such situations include that the interface only shows a rather small part of the database and does not provide much additional

information on the displayed movies. Movie information archives such as the *Internet Movie Database* (IMDb, <http://www.imdb.com>) cope with the latter issue by providing a structured, mostly textual overview of various details about a movie, not only including synopsis, runtime and actor information, but also information on writers, directors, filming location, and much more. Because this data is mostly displayed as text, we refer to this type of video database browsing as *text-based browsing*. While the presented information could potentially be very useful when searching for a movie to watch, the interface appears to be too overwhelming and, due to its heavy reliance on textual descriptions, less attractive and appealing.

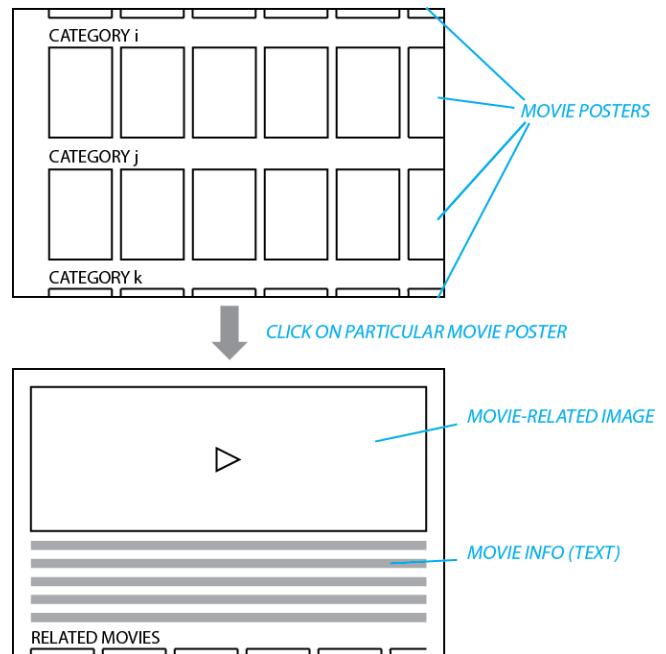


Figure 1. Illustration of common online movie database interfaces with a categorized overview of movie posters (top) and a detailed movie view (bottom). Categories can include, for example, new releases, personal recommendations, and genres, which are generally based on the user’s and other’s viewing behavior in the past. Movie details include some visuals representing it, accompanied by some text (short synopsis, runtime, etc.) and posters of a few related movies.

In this paper, we therefore introduce a new video database browsing concept called *visual navigation*, which aims at combining the best of both approaches; an interface that is as simple, intuitive, and appealing as visual browsing, yet also almost as powerful as text-based browsing. The key

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

TVX '17 Adjunct, June 14-16, 2017, Hilversum, Netherlands

© 2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-5023-5/17/06...\$15.00.

<http://dx.doi.org/10.1145/3084289.3089921>

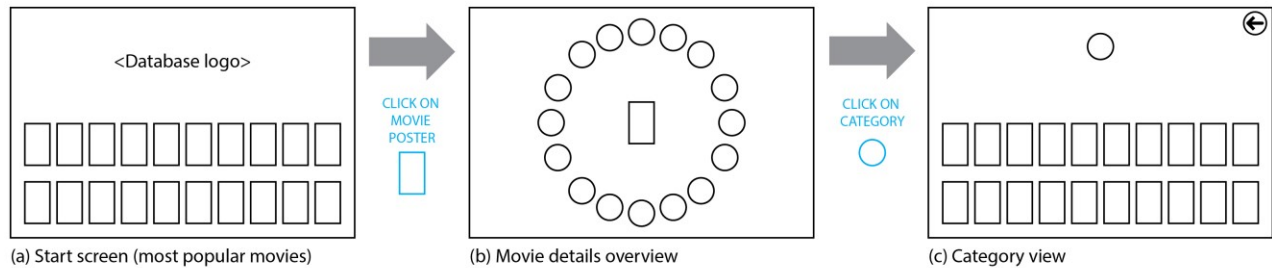


Figure 2. Illustration of the concept of *visual navigation* and screenshots of implementation used in the user study. Notice that visuals used in the test have been replaced here with copyright free material from Wikimedia.

idea is to visualize some of the information enlisted in text-based browsing approaches in an intuitive way, thus offering users the opportunity to easily explore related contexts; cf. Fig. 2.

While this idea of visual navigation seems simple, it is in no way obvious if such an interface would achieve the benefits we expect. Does it really provide a simplicity that is comparable to visual browsing interfaces? Can users understand and handle it? And does it really give you better access to the information displayed in text-based browsing approaches? Is that information helpful in finding better movies to watch? To get an initial understanding of the power of this concept of visual navigation, we present an informal pilot study addressing users' reaction to the idea, verifying if they can understand and handle the concept, and identifying possible advantages, issues, and pitfalls.

RELATED WORK

Common video database systems offer at least a simple search field enabling users to search for movie titles, actors, but also other things, such as directors and writers. Yet, in practice, most users generally restrict their search to the first two (titles and actors). Advanced searches on other movie characteristics, such as directors, filming locations, etc., could be very helpful in finding a movie, but are often not done, either because they are too complex to phrase or not supported by the system. Thus, providing an interface

that enables *intuitive* search based on such advanced movie characteristics is one of the key goals for our approach.

Video browsing is a topic that gained increasing attention in the video search and retrieval community in recent years. Good overviews of basic concepts and latest developments are provided by Schoeffmann et al. [5, 6]. Yet, most of this work addresses content-related search within video files, but lacks in supporting higher level characteristics, which we consider essential to create a more sophisticated video browsing experience. Thus, providing an interface that enables *powerful* higher-level concept browsing is another key goal for our approach.

In contrast to this work on content-based search within single videos, research related to exploring whole databases is often focused on recommendation systems, which are, for example, based on user profiles [7] or usage-driven [2]. Newer work also tries to accommodate for the fact that preferences can change and thus consider more context and user related issues (see [1, 2, 8], for example). While such systems work well if and only if the recommendation fits to a user's current need, this need cannot always be classified correctly (e.g., users themselves might not be sure or aware of it), it can vary, or it can have contradictory sub-aspects. In such situations, it is beneficial to complement such recommendation-based search with exploratory search.

Thus, providing the option to *actively explore* the video database is another key goal for us.

Active exploration is supported by Low et al.'s approach [4] where movies are represented and thus navigated using a metadata based clustering approach. Although quite related to our idea, the approach differs insofar as the reasons for these clusters (e.g., metadata such as genre, actors, director) are not visualized to the user. We claim that explicitly providing this information to a user can have a benefit not only in search performance, but can also make the search process more interesting and engaging. Particularly, we believe that the search for a movie to watch could and should itself be an engaging, entertaining, and fun experience. Thus, providing an *engaging* browsing experience is another key goal we are aiming for.

PROTOTYPE IMPLEMENTATION & STUDY SETUP

Based on the questions stated at the end of the first section and the criteria specified in the preceding one, we performed a pilot study to verify the potential of our idea for visual navigation. Particularly, we were interested in the following three issues:

1. Is our intuition about the lack of existing systems and the resulting specification of goals for our approach correct?
2. Is our approach as simple and intuitive as we assume? That is, can people understand and handle it easily?
3. Is this concept a valid alternative to visual and text-based browsing? What are advantages and disadvantages?

Prototype implementation

In this first step, we are interested in the applicability of our ideas for touchscreen operated tablets. Other devices, such as smartphones, TVs, or PCs will be addressed in our future research. The prototype used in the pilot study and depicted in Fig. 2 was implemented on an Android-based tablet (HTC Nexus 9) with an 8.9-inch screen (2048x1536 pixels resolution, 4:3 aspect ratio). Posters, other visuals, movie information and meta data were taken from *The Movie Database* (TMDb, <https://www.themoviedb.org>), a collaborative database for movies comparable to the commercial IMDb. We excluded movies of 2017 to avoid having movies in the test that have not been released yet.

The interface starts with a title page showing the twenty most popular movies in the database according to the provided popularity ranks (Fig. 2a). Selecting a movie by tapping its poster evokes the representation of this movie's *detail overview* (Fig. 2b). Here, the movie's poster is centrally shown in front of a movie-related background image. Movie-related meta data is represented around the poster in circular-shaped visualizations. Meta data includes genre, the three main actors, director, writer, main music composer, production company, language, and year. Notice that not all this data is available for each movie. Selecting a category by tapping it evokes what we call the *category view* (Fig. 2c). Here, the 20 most popular movies from that category are shown. Notice that for some specific

categories or some actors there might be fewer than 20 movies in the database. Users can then select one of the proposed movies or go back to the previous one using a back button at the top right of the screen.

Participants, study design, and procedure

We invited ten volunteers (8 males, 2 females, 18-23 years, average age 20.5 years) to participate in the pilot study. We focused on younger participants due to the intended target user groups, that is, frequent online media consumers and early adopters. Users were asked about their viewing behavior, with five indicating that they watch TV/movies "very often", three "often", and two "sometimes". Three mentioned to use Netflix or a comparable service "very often", three "often", one "barely", and three "never". Three said to use IMDb or a comparable service "often", two "sometimes", four "barely", and one "never". There was no correlation between these characteristics and the gathered data. Studies took place in a neutral room with participants seated comfortably at a table (cf. Fig. 3).

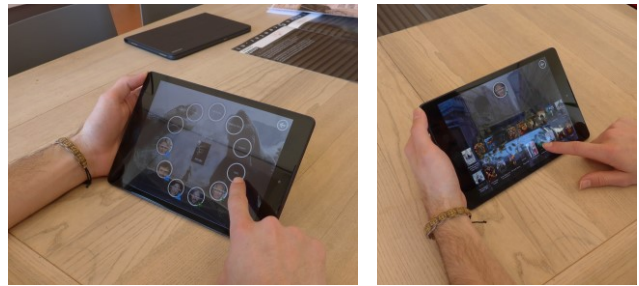


Figure 3. Illustration of the study conditions.

Despite its informal character, interviews with the subjects were scripted to avoid influencing their answers and achieve comparable feedback across participants. One subject was interviewed at a time. Duration was about 30 minutes per subject. Studies were split in three parts.

Part 1. To gain more insight into problems with standard interfaces (cf. first issue specified above), we started with an informal interview about users' viewing behavior and opinions on related interfaces. At the beginning, users were informed that:

"This is a study about a new alternative interface to access and browse online movie libraries."

After signing a consent form also informing subjects about the anonymous processing of their data, they provided demographic information and viewing behavior (cf. above), and were asked to explain what they liked and disliked about the interfaces of Netflix or similar services and IMDb or similar services. For the latter, subjects were also asked if they would consider it a useful complement to the first one if integrated into it. Screenshots of the interfaces were provided for subjects rather unfamiliar with these services. The interviewer wrote down all the subjects' answers.

Part 2. To verify if people can understand and handle our visual navigation concept (cf. second issue specified above)

we then let people explore the interface without further introduction. The interviewer only stated that:

“This is now a new, alternative interface for accessing and browsing online movie libraries. First, I want to see if you can figure out how it works on your own. Can you try to use it for, let’s say 2 minutes? Note that it’s not connected to the actual movie database yet, since this test is just about the new interface.”

After using the interface for two minutes without any interference of the interviewer, subjects had to rate five statements on a 5-point Likert scale from “strongly disagree” to “disagree” to “neutral” to “agree” to “strongly agree” (see results below for the concrete statements). Afterward, features not used or wrongly handled by the subjects were explained to them.

Part 3. To identify advantages, potential disadvantages, and possibly useful extensions (cf. third issue specified above), subjects then had to solve a concrete task, which was introduced by stating that:

“Now we want to test the interface for 5 minutes with a concrete task. Assume you are about to board a long flight, so you will be offline for a couple of hours. Hence, you want to download a couple of movies to watch when you are in the air. Now use the interface to search for, let’s say 3-5 movies that you would probably download. Once you found a candidate, just tell me and I write it down, since we haven’t implemented any bookmarking yet.”

After performing this task for five minutes without any interference of the interviewer other than writing down the subjects’ movie suggestions, they had to rate seven statements related to the interface using the same 5-point Likert scale as above. Following this, the session closed with an informal interview, where subjects were asked what they liked and disliked about the interface, how things they disliked could be improved, if the approach could be used as a replacement or complement to existing interfaces, and if they had any other comments or ideas for possible improvements. Finally, the interviewer discussed their provided ratings and other observations made during the tests with the subjects.

RESULTS & DISCUSSION

In the following, we summarize and discuss the results of the study. While we used the phrasings “Netflix / IMDb or similar services” during the interviews, we will subsequently use the terms “visual browsing” and “text-based browsing” describing the basic underlying concepts. Likewise, we will use “visual navigation” instead of “new interface”. Terms in brackets after a statement given below indicate how many subjects made this or a related comment.

Part 1 results (problems with standard approaches)

When asked about the positive aspects of visual browsing-style interfaces, subjects characterized them as visually

pleasing and “good looking” (6 subjects), and commented positively on the recommendations (2) and that the interface keeps track of what you watched (2). Most commonly stated negative aspects were that the category order seemed arbitrary and/or hard to find (4) and that searching takes long when one does not know what to watch (3). Other comments include that the amount of content is not enough (2) and that one sees the same movies often (2). These and additional individual statements seem to confirm our intuitive assumption about these types of interfaces stated in the first section.

While some subjects praised the additional functionality offered by text-based browsing interfaces and the provided structure, others described them as chaotic (4), stated that information is hard to find (2). While one subject said that “much of the information is useless”, half of them would see text-based browsing as a good complement when integrated into a visual browsing application (5). Thus, again, the subjects’ statements confirm our initial assumption that this data could potentially be very useful for video search if and only if it is represented in a way that is simple, structured, and easy to handle and understand.

Part 2 results (understanding and handling)

Table 1 illustrates the answers to the statements rated by the subjects. We see that people give high ratings for the intuitiveness and ease of use of visual navigation. Observations of the interviewer also suggest only minor issues with it that are mostly due to lack of experience; three subjects initially tried to swipe instead of clicking on categories or poster thumbnails, one started clicking on things too fast. (The transitions between the two views were animated.)

	Strongly disagree	Dis-agree	Neutral	Agree	Strongly agree
S1	0	0	0	4	6
S2	0	0	0	5	5
S3	0	0	0	5	5
S4	0	0	1	7	2
S5	0	2	2	4	2

S1: I understood the purpose of this application.

S2: The application worked intuitively.

S3: Navigating between movies was easy and straightforward.

S4: I felt in control of the application.

S5: I felt like all information that I needed was there.

Table 1. Number of subjects providing the related rating (column) for the given statement (row).

Despite the additionally displayed information compared to standard visual browsing-based interfaces, we can observe that subjects still felt that providing additional data would be good. When asked, most of them mentioned the information that is shown in the textual descriptions on the

details page for visual browsing (cf. Fig. 1, bottom), for example, run time and short synopsis. Such info could easily be integrated into our design using the empty spaces next to the visuals (cf. screenshots in Fig. 2). Yet, finding the right balance between how much and what kind of information should be displayed at all and in what way (textual or visual) is no trivial task and needs to be addressed in future research.

Part 3 results (advantages, disadvantages, ideas)

Statements in the questionnaire, illustrated in Table 2, cover different aspects. Subjects gave high ratings for ease of use (S1/S2), which is in line with the comments, ratings, and observations of part 2. Considering availability of information (S3), ratings are more neutral, which is again in line with part 2. It can be seen though that the tendency towards a lower rating is even stronger when faced with a concrete task. Again, this supports our claim that users appreciate or even demand more information when exploring movie databases. Keeping in mind that our approach provides much more context information about a movie than standard visual browsing interfaces, it is no surprise then, that most participants see visual navigation as a good complement (S5) or even replacement (S4) to those. A similarly strong trend cannot be observed for text-based browsing however (S6/S7). It is unclear though if this really suggests that our approach is considered to lack in power, or if subjects consider the lacking parts to be of lesser need when searching for videos.

	Strongly disagree	Dis-agree	Neutral	Agree	Strongly agree
S1	0	0	1	7	2
S2	0	0	2	6	2
S3	0	2	5	3	0
S4	0	0	3	4	3
S5	0	0	0	7	3
S6	0	1	4	2	3
S7	1	1	2	4	2

- S1: It was easy to find movies that I wanted to watch.
- S2: It was easy to navigate between the movies.
- S3: There was enough information for each movie.
- S4: I would use this application instead of Netflix for finding movies to watch.
- S5: I would use this application as a complement to Netflix for finding movies to watch.
- S6: I would use this application instead of IMDb for finding movies to watch.
- S7: I would use this application as a complement to IMDb for finding movies to watch.

Table 2. Number of subjects providing the related rating (column) for the given statement (row).

The subjects’ comments during the subsequent interview support these observations. Most described the interface as visually appealing and easy to use. Related phrases included “feels natural”, “well-designed”, “easy to use” (6 subjects). Some also appreciated that one can “build on previous search queries and don’t has to keep starting over” (5). Some mentioned that you get “good suggestions” (3) and that “lot of information (is) present” (2). On the negative side, only one subject stated that the “UI (is) not clear enough.” All other negative comments expressed the desire for additional functionality: a (categorized) search functionality (8), more information per movie, e.g., runtime, more actors (4), ratings / reviews (3), a home button (3), trailers (1), an overview of categories (1), and more detailed ones (1).

This, again, supports our claim that users prefer access to more information when searching for movies than usually provided by standard systems. Two subjects made comments during the test that are particularly interesting in this context: One said that “I never would’ve thought to search on production company, this is nice!” After clicking on the icon of an actress, another one asked: “Does she really play in all those movies?” These statements suggest that our approach does not only provide additional information to users that they are otherwise missing, but that it can also encourage them to search differently and help them to find movies that they would otherwise miss or ignore.

While one subject stated that “there is a good balance between the amount of information and easy browsing,” we do consider it a non-trivial and maybe the most important question for future research to find this perfect balance. Two subjects highlighted that it could happen that one “enters a loop of the same movies”. We often hear similar statements when people click on the “related movies” suggestions in visual browsing interfaces and the same movies are recommended repeatedly. Yet, it is certainly a risk in our approach as well. One might be able to resolve this issue though by finding the best information-versus-complexity balance as more navigation options will make this less likely.

CONCLUSION & FUTURE WORK

The design of better interfaces for online video archives is a problem of high relevance due to the increasing amount of data. Part 1 of our pilot study confirmed our initial, intuitive assumptions about common problems with existing interfaces, which mostly rely on the concept of visual browsing. Parts 2 and 3 confirmed that our alternative idea, entitled visual navigation, could provide a solution to deal with these problems: the subjects’ answers and comments suggest a need for providing more information and that visualization could be a good and intuitive way to do this. Yet, it also became clear that what to visualize and how is a critical and non-trivial question.

General feedback for the prototype implementation was overwhelmingly positive. One user even said this is “exactly what I want for Netflix.” The more critical comments suggest that this positive impression is not just due to the commonly observed “newness factor”, but based on educated reasoning. Also, we restricted the subjects to a group of early adopters and tech-savvy users. While this makes the statements more reliable, it does not allow us to generalize observations about intuitiveness and ease of use to the general population. This is in fact one of the most important and hardest challenges we are facing when designing interfaces for online movie databases: that they must be so intuitive and easy to handle that even totally inexperienced users can work with them (and enjoy doing this), while at the same time providing all the necessary power and functionality more experienced users are demanding for.

The positive results of our pilot study are based on feedback from the latter group, i.e., experienced users. Their encouraging comments and positive recommendations suggest several improvements, which we will address as a next step before testing a revised and enhanced design with broader and more heterogeneous user communities. Important questions to answer in this context include: What textual information should be added? What categories should be used and how should they be visualized? And, most importantly, what is the optimum tradeoff between providing additional information and not overwhelming the user or complicating the interaction? Likewise, we need to investigate what other functions (e.g., standard search, history and “go back” options) should be added and how they can best be integrated.

REFERENCES

1. Arapakis, I., Moshfeghi, Y., Joho, H., Ren, R., Hannah, D. and Jose, J.M., 2009, July. Enriching user profiling with affective features for the improvement of a multimodal recommender system. In *Proceedings of the ACM International Conference on Image and Video Retrieval* (p. 29). ACM.
2. Bollen, J., Nelson, M.L., Geisler, G. and Araujo, R., 2007. Usage derived recommendations for a video digital library. *Journal of network and computer applications*, 30(3), pp.1059-1083.
3. Hua, X.S. and Zhang, H.J., 2004, November. An attention-based decision fusion scheme for multimedia information retrieval. In *Pacific-Rim Conference on Multimedia* (pp. 1001-1010). Springer Berlin Heidelberg.
4. Low, T., Hentschel, C., Stober, S., Sack, H. and Nürnberger, A., 2017, January. Exploring Large Movie Collections: Comparing Visual Berrypicking and Traditional Browsing. In *International Conference on Multimedia Modeling* (pp. 198-208). Springer, Cham.
5. Schoeffmann, K., Hopfgartner, F., Marques, O., Boeszoermenyi, L. and Jose, J.M., 2010. Video browsing interfaces and applications: a review. *Journal of Photonics for Energy*, pp.018004-018004.
6. Schoeffmann, K., Hudelist, M.A. and Huber, J., 2015. Video interaction tools: a survey of recent work. *ACM Computing Surveys (CSUR)*, 48(1), p.14.
7. Shardanand, U. and Maes, P., 1995, May. Social information filtering: algorithms for automating “word of mouth”. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 210-217). ACM Press/Addison-Wesley Publishing Co.
8. Yang, B., Mei, T., Hua, X.S., Yang, L., Yang, S.Q. and Li, M., 2007, July. Online video recommendation based on multimodal fusion and relevance feedback. In *Proceedings of the 6th ACM international conference on Image and video retrieval* (pp. 73-80). ACM.