Scalable MovieBarcodes – An Exploratory Interface for the Analysis of Movies

Manuel Burghardt, Michael Kao & Niels-Oliver Walkowski
Computational Humanities Group, Leipzig University

ABSTRACT

In this article we present an exploratory interface for the analysis of movies. Movies are segmented into shots, which are in turn displayed as scalable MovieBarcodes, i.e. film scholars can zoom into the MovieBarcode representation to explore single chapters or scenes. The tool also provides a search function that can be used to filter shots according to characters or keywords, which are extracted automatically from subtitles and movie scripts. The filtered results are also displayed as interactive MovieBarcodes. Our tool can be used to aid film scholars during the research process of a movie analysis, as it provides new perspectives on a continuous, time-based medium.

Keywords: movie barcodes, scalable reading, visual movie analytics, exploratory interfaces, digital humanities.

Index Terms: H.5.m [Information Interfaces and Presentation (e.g., HCI)]: Miscellaneous;

1 INTRODUCTION

Since Roberto Busa's famous Index Thomisticus [4], the Digital Humanities have had a strong focus on text-based media. However, we are currently witnessing an increased interest for other types of media, most recently film and video [7, 18, 22]. Traditionally, empirical film studies have been working with shots as the smallest definable element that can be analyzed [21], i.e. they have been measuring and calculating the lengths and frequencies of different shots in movies [20, 27]. More recently, scholars have been thinking about other computable features of films [17], e.g. color variance, motion content, lighting key or the language of movies via subtitles [3, 5, 26]. While there are a plethora of different computational features available in film, it is still difficult for film scholars to investigate quantitative research questions, as the technical requirements to analyze the aforementioned features are still very high and typically require advanced knowledge in computer science. We believe that visualizations can be an important tool to make quantitative and qualitative aspects of movies available for analyses by a non-technical audience. More concretely, visualizations can aid film scholars to "reverse engineer" certain aspects of films and to reveal interesting patterns and trends at a larger scale [12].

Following this line of though, we propose an exploratory interface for a **visual movie analytics** approach [11]. The tool extends the visualization concept of **MovieBarcodes** [1, 3] by a **scalable reading** [14] approach, borrowed from the Digital Humanities.

burghardt@informatik.uni-leipzig.de
micha.kao@posteo.de

[1] **TRECVID**: https://trecvid.nist.gov/

Note: all URLs in this paper were last checked on July 9, 2018.

2 RELATED WORK

Numerous tools for the analysis and visualization of movie data have been proposed during the last years (for an overview see [23]. This is partly reflected by a dedicated track (TRECVID[1]) for video retrieval evaluation at the *Text Retrieval Conference* (TREC). While most of the existing tools and visualizations have been developed by the computer science community, there is also an increasing number of visualizations with a dedicated focus on Digital Humanities. Manovich [12] provides a set of tools that heavily relies on visualization as an exploratory interface to Dziga Vertov's movies. *ScriptThreads* [8] is another tool that can be used to visualize relations between the characters of a film. Other tools take into account the relation of dominant colors and sentiment information derived from movie subtitles [3, 6]. Another branch of related work builds on visual analytics for the investigation of video material in the Digital Humanities [10, 11].

The visualization proposed in this article can be classified as an application for browsing, retrieval and summarization of movies. The underlying design considerations and interaction concepts are described in the following paragraph.

3 DESIGN RATIONALES

The tool design is inspired by a previously conducted series of interviews with 4 students and 3 scholars of film and media studies. We asked them about typical scholarly practices and workflows for the analysis of movies, to gather requirements for a supporting tool. Most of the users reported that – among other factors, such as camera usage and montage techniques – the analysis of characters and dialogs are fundamental aspects of any scholarly film study, which was one of the driving aspects for the design of our tool. The tool is also influenced by the vast availability of subtitles and movie scripts, which can be used to extract character information and dialogs [5, 26].

The interviews also revealed that the repeated watching of a movie is a recurring theme in the scholarly process, as this is vital to establish, refine and ultimately answer a research question according to the hermeneutic circle approach [15]. While the usage of an exploratory interface obviously cannot substitute watching the movie itself, we still believe that it can support the initial process of generating and validating a research question by providing alternative perspectives on a movie. These different perspectives are realized by means of various search and filter functions as well as a scalable visualization concept.

4 FOUNDATIONS OF "SCALABLE MOVIEBARCODES"

The proposed tool builds on the visualization concept of *MovieBarcodes*, where each frame of a movie is skewed to 1px width. The condensed frames are then lined up in a barcode-like

representation, showing the most dominant colors of a movie and also representing the time axis of the movie from left to right (see Fig. 1).

Figure 1: MovieBarcode visualization of the movie "Eternal Sunshine of the Spotless Mind" (Michel Gondry, 2004).

While the basic idea of the MovieBarcode visualization may be traced back as early as 2001 [1], until today this visualization format is very popular among film enthusiasts in the social media context. Most notably, there is a lively community on *Tumblr* that has coined the term "MovieBarcode" for this type of visualization and that, so far, has created an impressive collection of approx. 3,000 MovieBarcodes[2]. This collection has also been used to create a color-based IR system [2], allowing scholars to search for movies according to dominant colors and other metadata. An adaption of the MovieBarcode visualization can be found in Frederic Brodbeck's cinemetrics project[3].

MovieBarcodes can be considered a color-based fingerprint of a movie that enables film scholars to get a basic overview of a movie and to compare a large number of movies on a very generic level. In literary studies, "distant reading" [13] has been suggested as a term for this mode of analysis. While literary scholars are typically known for close reading, i.e. the hermeneutical, in-depth analysis of a few selected texts, distant reading makes visible hidden connections and overarching patterns in larger collections of texts. While there are strong arguments for both, close and distant reading [9], a hybrid approach that allows scholars to switch between both modes of reading seems most promising. Accordingly, Martin Müller has suggested a "scalable reading" approach, allowing users to zoom in and out on an interactive text [14]. Müllers idea is highly reminiscent of existing design guidelines for information visualization ("Overview first, zoom and filter, then details-on-demand"; [24]). In the Digital Humanities context, interfaces that enable scalable views with different levels of detail on texts and other types of media have been termed rich-prospect browsers [19].

The following paragraph illustrates how we enhanced the static MovieBarcode visualization by a scalable reading interaction concept.

5 Visualization and Interactions

Assuming that shots are the smallest meaningful unit of interpretation [21], they are also used as the basis of our scalable MovieBarcodes. Each shot is represented by one keyframe in the visualization (see Fig. 2). The default view is 1% zoom, i.e. the whole movie is represented as a traditional MovieBarcode that gives a broad overview of the movie. From here, users may change the scale of the visualization by means of a slider that can be used to zoom in or zoom out. In addition, it is possible to navigate the MovieBarcode by means of the DVD chapters.

While MovieBarcodes typically provide an overview of the most dominant colors, we added additional information about the characters and dialogs, to provide a more comprehensive summary of the movie. Below each of the keyframes, dialogs are visualized as a colored rectangle, where colors are used to distinguish the characters. Each rectangle represents an utterance from the movie's subtitles. By adding this information to the MovieBarcodes, film scholars can look at films from a distance, to identify shots with more or less dialogs, or even search for shots with dialogs from a specific character. To aid these visual cues on the distribution of character speeches in the movie, we also provide a search bar that can be used to search for one or more characters that appear in the dialogs. In addition, shots may be filtered according to keywords that appear in the corresponding dialogs. The results are also displayed as scalable MovieBarcodes, with the option to show or hide the filtered shots, i.e. shots that are not in the result set. Film scholars may click at a keyframe at any time to get a detailed view (close reading) of the shot and the corresponding dialogs (see Fig. 3).[4]

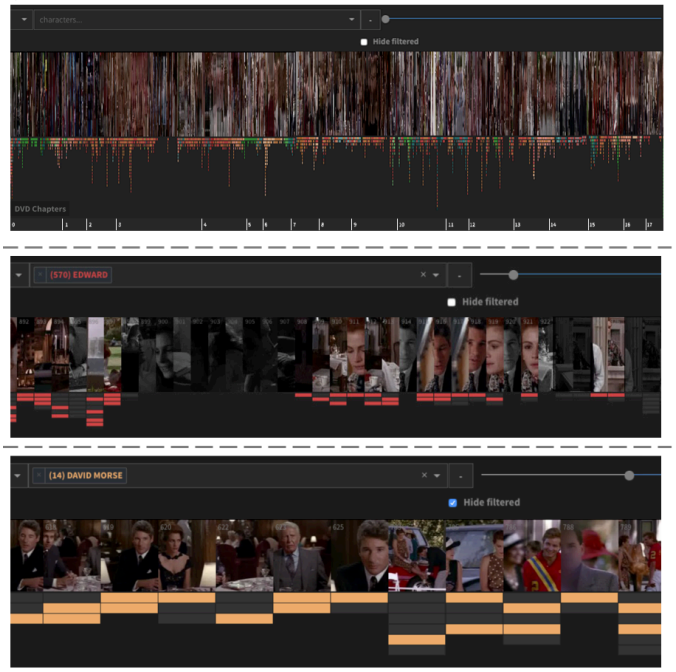

Figure 2: The three MovieBarcodes illustrate different levels of zoom for shots from "Pretty Woman" (Garry Marshall, 1990). The top image shows a condensed MovieBarcode summary of the whole movie. The middle image shows all shots that contain the character "Edward" in the MovieBarcode, any other shots are displayed in greyscale. The bottom image shows all shots that contain the character "David Morse", any other shots are excluded from the MovieBarcode ("hide filtered" option activated).

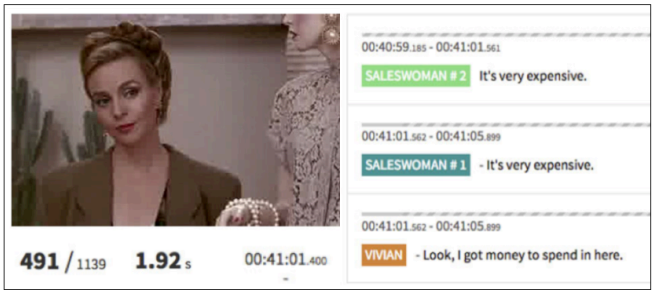

Figure 3: Detailed view of shot 491 in "Pretty Woman" (Garry Marshall, 1990)

[2] **MovieBarcode Tumblr**: http://moviebarcode.tumblr.com/movie-index

[3] **Cinemetrics**: http://cinemetrics.fredericbrodbeck.de/

[4] A **video demo** is available at https://vimeo.com/209187351

6 Architecture and Components

The interactive visualization was implemented as a web application that has two main components (see Fig. 4): The **analyzer** component is written in Python and heavily relies on the OpenCV library[5]. The analyzer is used for processing and segmenting the video data and DVD chapter information. For the shot boundary detection, we implemented both, an approach based on color histograms and an approach that relies on edge detection [25]. While both methods yield decent results, the histogram-based algorithm is way more efficient with approx. 15 minutes of calculation time as compared to 10-20 hours for the edge detection approach. The analyzer is also used for text mining of the dialogs. We use subtitles[6] to extract time stamps ("Which shot does the dialog belong to?"), and movie scripts[7] to extract character information ("Who speaks?"). For the alignment of both, timestamps and speaker information, we implemented an algorithm that relies on the dynamic time warping technique [5, 16, 26].

The actual **visualization** is realized as a viewer component that communicates with the backend via a simple REST API. The backend is connected to a database that contains all the necessary information about shots, chapters, dialogs and characters. An asset store provides the keyframes that are used to dynamically generate the visualizations in the viewer component. Both, the analyzer (together with a basic documentation of its features and their technical realization) and the visualization framework are available via GitHub[8].

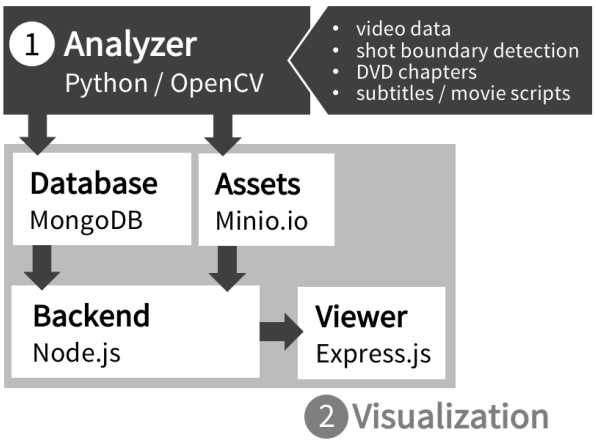

Figure 4: System architecture and components.

7 Conclusion

In this paper we propose a tool that adopts the scalable reading approach known from literary studies for the analysis of movies. The tool extends the visualization paradigm of colored MovieBarcodes by making them interactively scalable and by adding additional information about characters and dialogs. The exploratory interface allows scholars to compare different movies with respect to dominant colors and dialog structures. On a rather generic level of analysis (distant reading), scholars may explore structural patterns for a corpus of movies. Film scholars could for instance compare successful Hollywood movies, to explore whether the usage of bright or dark colors or the proportion of dialogs across those movies are rather similar or totally arbitrary. For interesting movies that break an identified pattern, scholars may then zoom into the MovieBarcode, to pursue a close reading approach and to investigate the function of specific characters or the keywords throughout the movie, which might be another indicator for its success.

Until now, the tool may still be considered a prototype that was designed to meet some fundamental requirements of scholarly movie analysis. We are currently working together with film scholars to gather concrete use cases for the application of the Scalable MovieBarcodes tool for different scholarly tasks and research questions and to derive best practices for conducting computer-supported analyses of movies.

References

[1] Barbieri, M., Mekenkamp, G., Ceccarelli, M., & Nesvadba, J. (2001). The color browser: a content driven linear video browsing tool. In Multimedia and Expo Conference, IEEE, pp. 627–630.

[2] Burghardt, M., Hafner, K., Edel, L., Kenaan, S., & Wolff, C. (2017). An Information System for the Analysis of Color Distributions in MovieBarcodes. In Proceedings of the 15th International Symposium of Information Science (ISI 2017).

[3] Burghardt, M., Kao, M., & Wolff, C. (2016). Beyond Shot Lengths – Using Language Data and Color Information as Additional Parameters for Quantitative Movie Analysis. In Book of Abstracts of the International Digital Humanities Conference (DH), Krakow.

[4] Busa, R. A. (1980). The Annals of Humanities Computing: The Index Thomisticus. Computers and the Humanities, 14(2), 83–90.

[5] Everingham, M. R., Sivic, J., & Zisserman, A. (2006). Hello! My name is... Buffy – Automatic Naming of Characters in TV Video. Proceedings of the British Machine Vision Conference 2006, 92.1-92.10. http://doi.org/10.5244/C.20.92

[6] Hohman, F., Soni, S., Stewart, I., & Stasko, J. (2017). A Viz of Ice and Fire: Exploring Entertainment Video Using Color and Dialogue. Proceedings of the 2nd Workshop on Visualization for the Digital Humanities (VIS4DH), Phoenix, Arizona.

[7] Hoyt, E. & Acland, C. R. (eds.) (2016). The Arclight Guidebook to Media History and the Digital Humanities (2016). University of Sussex. Online version: http://projectarclight.org/wp-content/uploads/ArclightGuidebook.pdf

[8] Hoyt, Eric, Ponot, Kevin, & Roy, Carrie (2014). Visualizing and Analyzing the Hollywood Screenplay with ScripThreads. In: Digital Humanities Quarterly 8(4)

[9] Jänicke, S., Franzini, G., Cheema, M. F., & Scheuermann, G. (2015). On Close and Distant Reading in Digital Humanities : A Survey and Future Challenges. Eurographics Conference on Visualization (EuroVis) (2015), 1–21.

[10] John, M., Kurzhals, K., Koch, S. & Weiskopf, D. (2017). A Visual Analytics Approach for Semantic Multi-Video Annotation. Proceedings of the 2nd Workshop on Visualization for the Digital Humanities (VIS4DH), Phoenix, Arizona.

[11] Kurzhals, K., John, M., Heimerl, F., Kuznecov, P., & Weiskopf, D. (2016). Visual Movie Analytics. In IEEE Transactions on Multimedia 18(11), 2149-2160.

[12] Manovich, L. (2013). Visualizing Vertov. Russian Journal of Communication, 5(1), 44–55.

[13] Moretti, F. (2000). Conjectures on world literature. New Left Review, (Jan / Feb), 54–68.

[14] Müller, Martin (2012). Blog on Scalable Reading – Dedicated to DATA: digitally assisted text analysis. Online: https://scalablereading.northwestern.edu/?page_id=22

[5] **OpenCV**: https://opencv.org/

[6] **OpenSubtitles**: https://www.opensubtitles.org/

[7] **IMSDb**: http://www.imsdb.com/

[8] **Analyzer**: https://github.com/Miiha/FilmAnalyzerKit; **Visualization**: https://github.com/Miiha/FilmBrowser

[15] Palmer, R. (1969). "Hermeneutics: Interpretation Theory in Schleiermacher, Dilthey, Heidegger, and Gadamer". Evanston: Northwestern University Press.
[16] Pramod, S., Jawahar, C. V, & Zisserman, A. (2009). Subtitle-free Movie to Script Alignment. Movie, 1, 1–11.
[17] Rasheed, Z., Sheikh, Y., & Shah, M. (2005). On the use of computable features for film classification. IEEE Transactions on Circuits and Systems for Video Technology, 15(1), 52–63.
[18] Ross, M., Grauer, M. & Freisleben (eds.) (2009). Digital Tools in Media Studies. Analysis and Research – An Overview. transcript.
[19] Ruecker, S., Radzikowska, M., & Sinclair, S. (2011). Visual Interface Design for Digital Cultural Heritage. Farnham et al.: Ashgate Publishing.
[20] Salt, B. (1974). Statistical Style Analysis of Motion Pictures. In: Film Quarterly, 28(1), 13-22.
[21] Salt, B. (2006). Moving into Pictures. More on Film History, Style, and Analysis. London: Starword Publishing.
[22] Sayers, J. (ed.) (2018). The Routledge Companion to Media Studies and Digital Humanities. Routledge.
[23] Schoeffmann, K., Hudelist, M. A., & Huber, J. (2015). Video interaction tools: A survey of recent work. ACM Computing Surveys (CSUR) 48(1), 1–34.
[24] Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. Proceedings of the IEEE Symposium on Visual Languages (pp. 336-343).
[25] Smeaton, A. F., Over, P., & Doherty, A. R. (2010). Video shot boundary detection: Seven years of TRECVid activity. Computer Vision and Image Understanding, 114(4), 411– 418. http://doi.org/10.1016/j.cviu.2009.03.011
[26] Cour, T., Jordan, C., Miltsakaki, E. & Taskar, B. (2008). “Movie/script: Alignment and parsing of video and text transcription,” in Proc. Comput. Vis., 158–171.
[27] Tsivian, Y. (2009). Cinemetrics, Part of the Humanities' Cyberinfrastructure. In Ross, M., Grauer, M. & Freisleben (eds.) (2009). Digital Tools in Media Studies. Analysis and Research – An Overview. transcript.