

Methods and Advanced Tools for the Analysis of Film Colors in Digital Humanities

Barbara Flueckiger <baflueckiger_at_gmail_dot_com>, University of Zurich
Gaudenz Halter <N/A>, University of Zurich

Abstract

Colors are one of the most difficult stylistic elements of film to analyze, but — as this paper elaborates — a most rewarding one too. A long-neglected topic in film studies, film colors have gained increasing attention over the last decade. With the development of database-driven analysis, deep-learning tools, and a large range of visualization methods the research project ERC Advanced Grant *FilmColors* set out to provide a more comprehensive approach to analyzing the manifold aspects of color in film. This paper focuses on a set of strong theoretical and analytical concepts of film colors — including human interpretation — that connect the stylistic, expressive, and narrative dimensions with the development and evaluation of digital methods. A corpus of more than 400 films have been analyzed with a computer-assisted workflow that has been integrated into the video annotation and analysis software VIAN since 2017. VIAN is connected to the online platform VIAN WebApp for the evaluation of results, queries, and visualizations on segment, film and corpus level. Compared to traditional, mostly language-dominated approaches to the aesthetics, technology, and narratology of film colors, the digital humanities tools turn evidence created through the mapping of results into an instantly accessible array of visual representations. By relating detailed human annotation and interpretation to these visual representations, the integrated workflow consisting of the VIAN visual analysis software in combination with the crowdsourcing portal VIAN WebApp has created a comprehensive ecosystem for the investigation of film aesthetics and narration. It thus significantly extends established methods in film studies.

1 Introduction

Film colors are one of the most difficult aspects for the analysis of film style, but — as this paper will elaborate — also a most rewarding one. A long neglected topic in film studies, film colors have gained increasing attention during the last decade. Following the advent of digital methods in recent years, in-depth studies about the history, uses, underpinning concepts and their theoretical and epistemological reflection in digital humanities have been published by Olivia Kristina Stutz (2016), Adelheid Heftberger (2016, 2018), Christian Gosvig Olesen (2017). Increasingly there is a discourse around the development of digital approaches, methods, and tools for film analysis. Stutz (2016) and Olesen (2017) pay specific attention to the question of color analyses. Kevin Ferguson (2013 and 2015), Lev Manovich (2013 and 2015) and Everardo Reyes-García (2014 and 2017) have developed specific visualization methods for art works, paintings and film in particular. Pause and Walkowski's work on computer-assisted color analysis is drawing on our own work [Pasue and Walkowski 2018].

Traditional analyses of film colors were mostly based on verbal description. They showed a tendency toward hermeneutical interpretation while aesthetic and affective dimensions were often neglected.

With the development of database-driven analysis, deep-learning tools and a large range of visualizations the research project ERC Advanced Grant *FilmColors* (see Acknowledgements) aims at providing a more comprehensive approach to analyze the manifold aspects of color in film.

Therefore the central argument of this paper focuses on the combination of a set of strong theoretical and analytical concepts including human interpretation that connects the various instances of film colors' stylistic, expressive and narrative dimensions to the development and evaluation of digital methods.

It is a widespread misconception that digital tools generate meaningful results in an automated fashion. Theoretically sound reasoning and the constant reflection of visualization methods, their epistemological and perceptual underpinnings is a mandatory requirement that must govern any interdisciplinary collaboration between film studies and computer science, see our previous

1

2

3

4

5

papers for a more extended discussion of these prerequisites and their connection to experimental aesthetics [Flueckiger 2011] [Flueckiger 2017] [Flueckiger and Halter 2018]. By contrast to the previous papers, this article intends to provide insights into the many methods, obstacles, advances, problems and lessons learned during the collaborative development of the tools.

Integral part of the research projects is the *Timeline of Historical Film Colors* (<https://filmcolors.org>) — an interactive, comprehensive web resource on film colors that has been created and curated by Barbara Flueckiger since 2012 [Flueckiger 2012].

2 Database-driven Analysis, Analytical Concepts and Evaluation

Overarching goal of the research project's interdisciplinary perspective is the investigation of the relationship between the aesthetics and technology of film colors. To this end, a large group of more than 400 films produced between 1895 and 1995 were analyzed in a highly detailed way with a computer-assisted approach. It combined temporal segmentation by the video annotation tool ELAN (first released in 2002) with a database-driven protocol consisting of a controlled vocabulary of around 1.200 theoretical and analytical concepts for the annotation of each segment. A network of custom-made relational databases for the analysis, filmographic data, glossary and evaluation of results (see Figure 7) was programmed in FileMaker to a large extent by the PI herself with input from her team. Despite the fact that FileMaker has its weaknesses and limitations in terms of programmability and flexibility the self-sufficient adaptation and development of the databases according to the evolving requirements of the analyses remained the most significant advantage throughout the project. Increasingly, the databases were linked with each other by relational connections to deliver meaningful results and to provide instant access to a variety of evaluation methods of the data gathered (see Section 3).

The team distinguished several levels of analysis, from screenshots to temporal segments of individual films (“micro level”), individual films as a whole (“meso level”), and over the whole corpus or selected sub-corpora (“macro level”), see Olesen (2016).

Filmographic data has been collected to represent the whole corpus of films, to define corpus selection and assignment to individual researchers and to keep track of the processing state. Corpora were selected based on research into monographs and articles on film colors, with each of the PhD candidates' setting their own focus in the three periods 1895 to 1930, 1930 to 1955, and 1955 to 1995. Guiding principles for selection criteria aim at a comparison between canonical works, famous for their elaborate or bold color design with lesser known works to form sub-groups for specific film color processes, genres, for individual filmmakers, cinematographers, color consultants, or countries. In line with historical poetics [Bordwell 1989] and neo-formalist analysis established by the Wisconsin school of Kristin Thompson and David Bordwell [Thompson 1988] [Bordwell et al. 1985], the corpora should enable the *diachronic analysis* of personal styles, institutional contexts — for instance changing professional norms, cultural preferences, notions of taste — or technological innovation over time, but also a *synchronic comparison* between these different instances at a given period.

Stock identification, research into technical innovation, detailed information about color processes applied to each film play an important part for a better understanding of these connections and allow to circumvent misconceptions present in previous literature. These research methods are completed by scientific measurements of film stocks' spectral characteristics to enable improved methods for the digitization and restoration of analog film colors. Such a comprehensive approach that connects insights into aesthetic developments with a deep investigation into technological innovation has been called a “technobole approach” by Frank Beau (2002). Contrary to technical determinism the “technobole approach” that stays at the center of our method is paying attention to epistemological, institutional, social, cultural and economic factors that govern technical advances and how technology in turn feeds back into culture and society. This cultural perspective is investigated thoroughly in the PI's second research project *Film Colors. Technologies, Cultures, Institutions*.

A three-level model that complies with recommended metadata schemes established for film archives by standardization entities such as FIAF or filmstandards.org has been implemented into the corpus database by team member Joëlle Kost. It allows the allocation of individual tokens of a single film, such as DVDs, Blu-ray or various historical film prints and negatives inspected and documented in film archives in Europe, Japan and the United States and assigns a specific tri-partite item ID to each one of them. From the start, this database was hosted on a FileMaker server provided by the University of Zurich.

All the films chosen for analysis were digitized and then temporally segmented with the video annotation tool ELAN (see Section 4 for approach and description). The resulting information — mostly time codes and numbering of the segments plus optionally basic descriptions according to a template — was then exported to an analysis database for close reading, which in turn was connected to the corpus DB for filmographic data, based on the item ID.

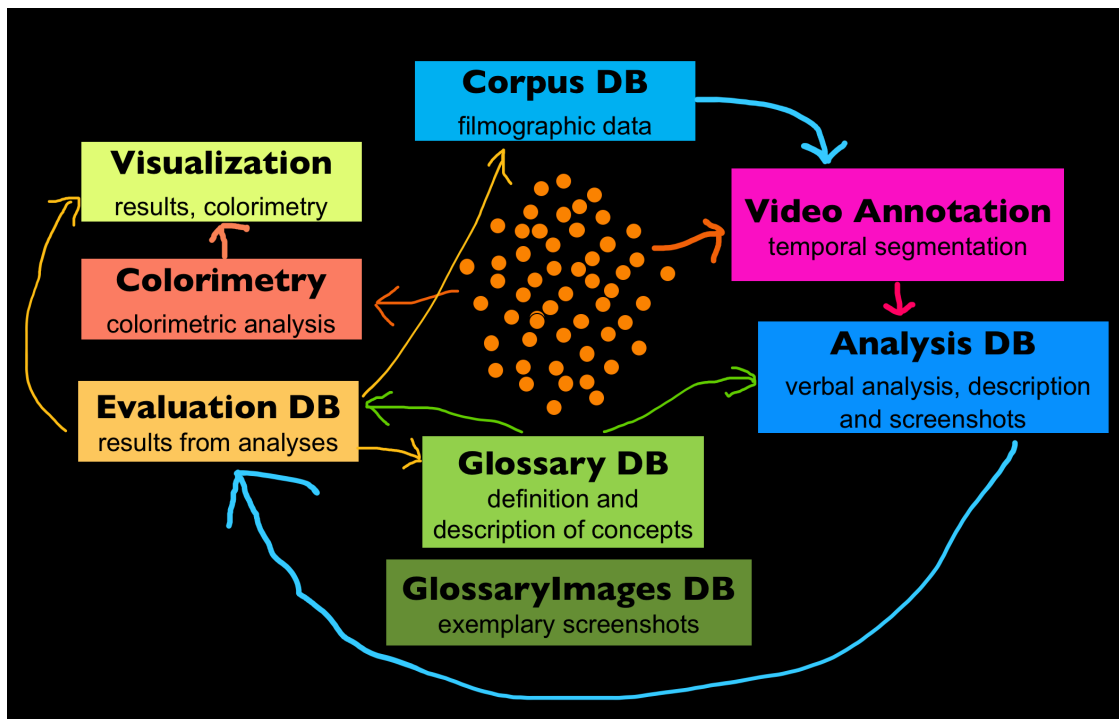


Figure 1. Sketch of the analysis and evaluation workflow with the database architecture.

Theoretical and analytical concepts were mostly elaborated before the start of the project, during the PI's teaching and research activities in the field of color theory, film aesthetics, semantics and narration — with a focus on film colors in the last decade. Accordingly, they were already part of the research proposal. Submitted to the ERC. During half a year of coaching and training at the beginning of the project, the team members were introduced into the concepts and were given room for extended discussions. Some concepts evolved bottom-up during the analyses and were contributed by team members based on their observations. For instance, a catalogue of basic terms for characters' affective or emotional states were part of the initial glossary, but continuously extended by postdoc researcher Bregt Lameris who focuses on the relationship between film colors and affects. Motifs and themes were also evolving bottom-up for certain standard situations — for instance ceremonies, show numbers, metamorphoses — or topics such as exoticism, architecture, self-reference etc. They were organized in a keyword database connected to the analysis DB.

Eleven different registers contained a taxonomy with classes of analytical concepts, ranging from verbal descriptions of colors, color contrasts, image composition, depth-of-field, lighting, textures and patterns, surface properties and materials of characters, objects or environments in the diegesis including their tactile properties, to the materiality of films analyzed with the concept of *faktura*, plus movements of camera, characters, objects, and lighting.

For each temporal segment of the films — usually between 50 and 70 segments per film — the team went through the whole range of analytical concepts in the analysis DB and added up to 32 relevant screenshots into media containers.

All the theoretical and analytical concepts of the controlled vocabulary — more than 1.200 including hues — were continuously defined in a glossary database with references to sources, if available, and illustrated with exemplary screenshots gathered during the analyses.

From the start we were thinking about how these concepts could be processed with advanced tools for automatic data extraction, and this question guides the development of deep learning tools to this day (see Sections 4 to 9). It goes without saying that not all of them can easily be solved with the current state-of-art and limited resources even within an ERC research project of this scope. On the other hand — as stated above — it is the central credo of the project's comprehensive approach that it aims at a qualitative analysis that focuses on the contextualization of observations by human intervention and interpretation. By their very nature, automatic approaches to image processing are not able to identify subtle details and idiosyncrasies, for instance that curtains moving slightly in the wind might be a metaphor for the heroine's inner turmoil as in the Japanese film *Jigokumon*.

Narratologic concepts are especially challenging for automatic assessment. Point-of-view structures that operate with the concept of *focalization* ([Genette 1972] [Genette 1983] to differentiate between instances of narration, so-called focalizers or filters, are

possibly very hard to identify for non-human observers, but they are very important for the investigation of film colors, including characters' mental states in dreams or hallucinations, alignment with characters [Smith 1995], temporal organization of the narration such as flashbacks, summaries, mise-en-abyme, parallel montage or montage sequences, non-narrative situations, turning points, suspense and foreshadowing etc.

Similar challenges are in operation for phenomena of higher order semantics. By higher order semantics we understand all forms of modification of meaning that are established through intra-textual relationships, intertextual and inter-medial references or cultural uses, such as cultural contexts, milieu, taste, sociopolitical markers, stereotypes, genre conventions, character relationships, race, gender, symbols, signals, metaphors and isotopies / rhyming. Intertextual and inter-medial references for connections to other films, media or art works through pastiche, allusion, citation, irony, parody etc. [Genette 1992] [Jameson 1991] [Dyer 2006].

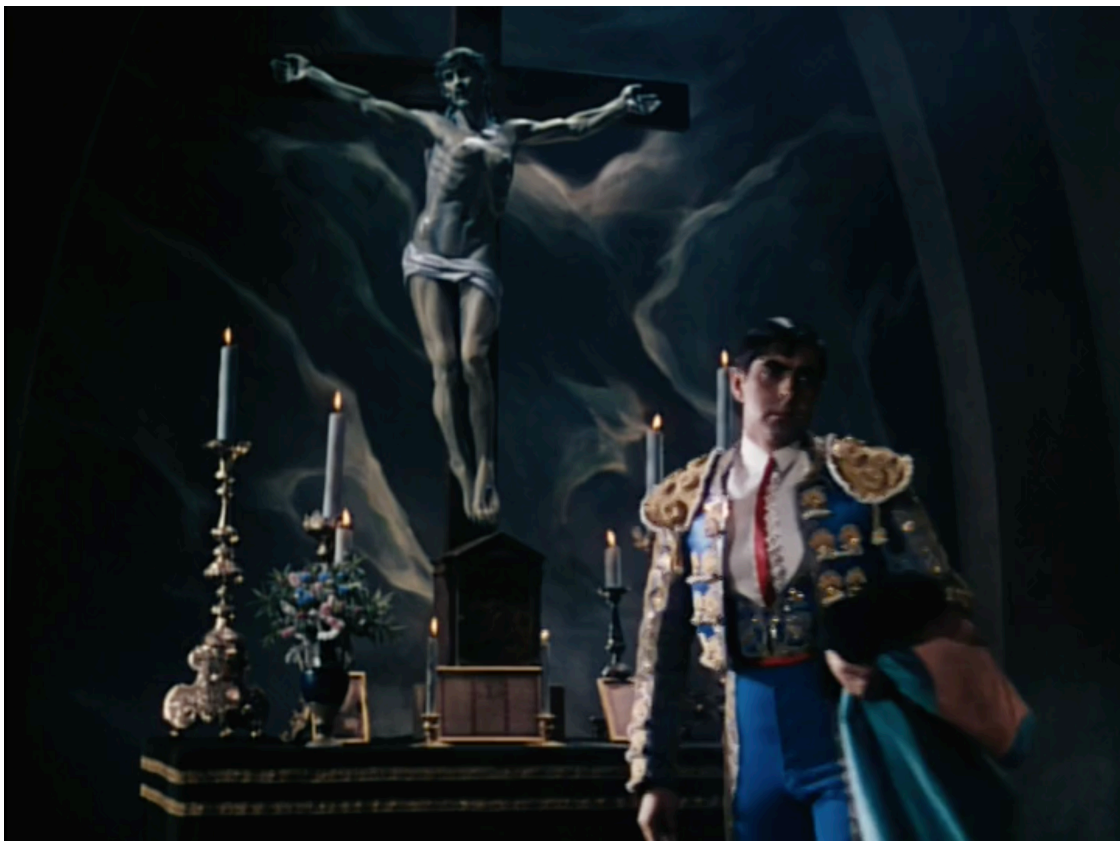


Figure 2. Higher order semantics in color design: pastiche in *Blood and Sand* (USA 1941, Rouben Mamoulian)(left) and sociopolitical markers that indicate status and period in *The Private Lives of Elizabeth and Essex* (USA 1939, Michael Curtiz) (right).

Emotions and affects relate to characters' inner states — joy, sadness, anger, hate, disgust — or emotional relationships such as love, conflict, sex, but also for cross-modal relationships of visual representations to smell, taste or touch. In addition, there are

basic theoretical concepts for emotional and affective responses such as direct affect [Plantinga 2009], contagion, artefact emotion, mimicry, plus aesthetic categories that address the senses such as excess [Thompson 1986], artefact emotion [Tan 1996], atmosphere, *Stimmung* or mood.



Figure 3. Excess in *King of Jazz* (USA 1930, John Murray Anderson; Pál Fejös) (above) and *Die bitteren Tränen der Petra von Kant* (GER 1972, Rainer Werner Fassbinder) (below).

In the domain of color identification, however, computer-based approaches are superior to human observation, even if it is necessary to stress the fact that the results of these analyses also need human interpretation.

In our analysis DB, colors were verbally described as dominant hues for the entire scene, female and male protagonists and supporting characters, background and foreground, inter-titles and letters, including general observations on saturation, lightness etc. It is obvious that such verbal descriptions are very limited, they do not take into account the subtle shades of each color of a certain range as various types of hues, levels of saturation or brightness.

By color schemes — often called *color palettes* — we denote the overall distribution of color in an image or in a temporal segment according to the variations of hues, saturation, warmth or lightness, such as monochrome restrictive, muted, gaudy, saturated etc. Types can occur simultaneously, for instance a monochrome color scheme can also be warm or saturated. As we will elaborate in a later section (see Section 7) our methods for the identification of color schemes with deep learning tools are both superior, more refined than verbal descriptions and yield highly significant results, if, again, they are connected to the concepts elaborated above.

Color contrasts refer to an established set defined by artist and scholar Johannes Itten to describe specific relationships of color harmonies, color “chords,” or spatial distribution, again correlated to the dimensions of hue, saturation and lightness as organized in Itten’s “Farbstern” (*color star*) [Itten 1970]. For instance, the most ubiquitous color contrast in the recent decade has been the orange–teal combination that is both a cold–warm and complementary contrast plus it contains a light–dark variation since yellow is perceived as brighter than blue.

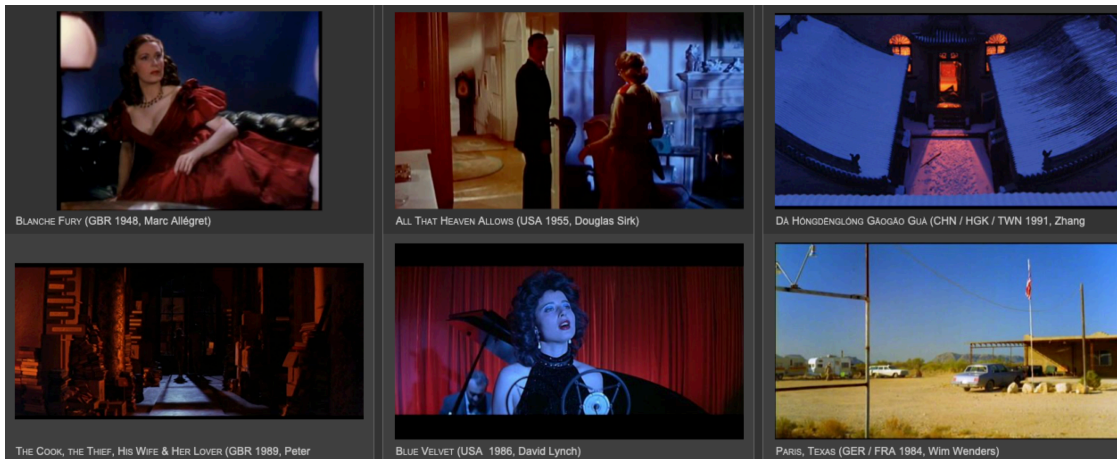


Figure 4. Some exemplary screenshots for the concept “cold-warm contrast” in the glossary DB.

The identification of color contrasts and color schemes has long been a field of computer analysis. As we will discuss (see Section 7), however, most of these approaches do not comply with the demands of aesthetic analysis in terms of differentiation, flexibility and subtlety. Some of them were established for normative purposes, to identify ‘good’ uses of color harmonies for design, or to give a rough, albeit pleasing visualization for film geeks, such as MovieBarcodes.

2.1 Problems

Consistently the biggest challenge was the level of complexity for all the team members working on the film analyses. Overall the process was perceived as extremely time consuming. Following a first evaluation after several months into the project, the concepts were separated into the most relevant ones vs. the rarer ones. Team members also showed difficulties to work on such extended catalogues of concepts that were only randomly ordered by relevance. Therefore they devised ordered lists sorted according to thematic coherence, which helped finding the checkbox in a more intuitive way. Finally we ended up establishing individual layouts for each team member to take their personal focus into account. However, this approach resulted in a much more complex database architecture to collect and evaluate all the data.

2.2 Lessons learned

Informed by the constantly evolving workflow and database architecture the crowdsourcing platform for external users [Flueckiger and Halter 2018] [Halter et al. 2019] has been developed in a modular fashion, again by Gaudenz Halter with support by Silas Weber. First, concepts are being sorted according to inner relationships, for instance positive affects related to joy vs. negative affects related to depression or aggression as elaborated and refined during the development of the workflow and following the final evaluation. Second, levels of significance of concepts and levels of complexity have been established for each area of analysis, as for instance lighting or image composition. This modular design will give users the possibility to select from a menu of

concepts not only the topics they are interested in, for instance color contrasts or costume design, but also the level of complexity for each of those modules individually so that the controlled vocabulary matches best their research interest.

3 Evaluation of the Data Gathered

Resulting from the manual analysis was a massive quantity of data and screenshots, amounting to more than 17.000 segments with about 170.000 screenshots assembled in a master analysis DB and more than half a million of summations gathered in an evaluation database. This evaluation database has been connected to the glossary DB and the corpus DB, based on the glossary ID and the Item ID (see Figure 1). With these identifiers we were able to then display the results directly in the glossary DB and corpus DB respectively by portals and scripts, which turned out to be the biggest advantage of the relational database architecture. As will be discussed in a later paragraph (see Section 3.1) there were also many problems and challenges to master.

In principle the results can be accessed by three ways through the FileMaker architecture and in many additional, more complex ways through the analysis and annotation tool VIAN and the online platform VIAN WebApp. The VIAN WebApp is a crowdsourcing portal that currently contains all the more than 400 analyzed films for evaluation and visualizations on segment, film and corpus level [Flueckiger and Halter 2018] [Halter et al. 2019]. In the future, external users will have the possibility to commit their own VIAN projects. Since the database for the VIAN WebApp hosts both qualitative and numeric color information, the developers decided on a diploid database architecture for the online platform. Most of the data is hosted on a Postgres SQL database; for fast querying and processing numeric information they use a HDF5 file structure (see Halter et al. 2019). Data are processed by cloud computing on Microsoft Azure.

The offline analysis VIAN is integrated into an ecosystem (see Figure 5). Individual projects are uploaded to the online platform VIAN WebApp. In return, users can download projects from the VIAN WebApp to adjust it to their own interests. In addition, the VIAN WebApp connects projects to the corresponding galleries from the *Timeline of Historical Film Colors*. Finally, the ColorMania app became an extension for visitors of the *Color Mania* exhibition at Fotomuseum Winterthur (see Figure 5).

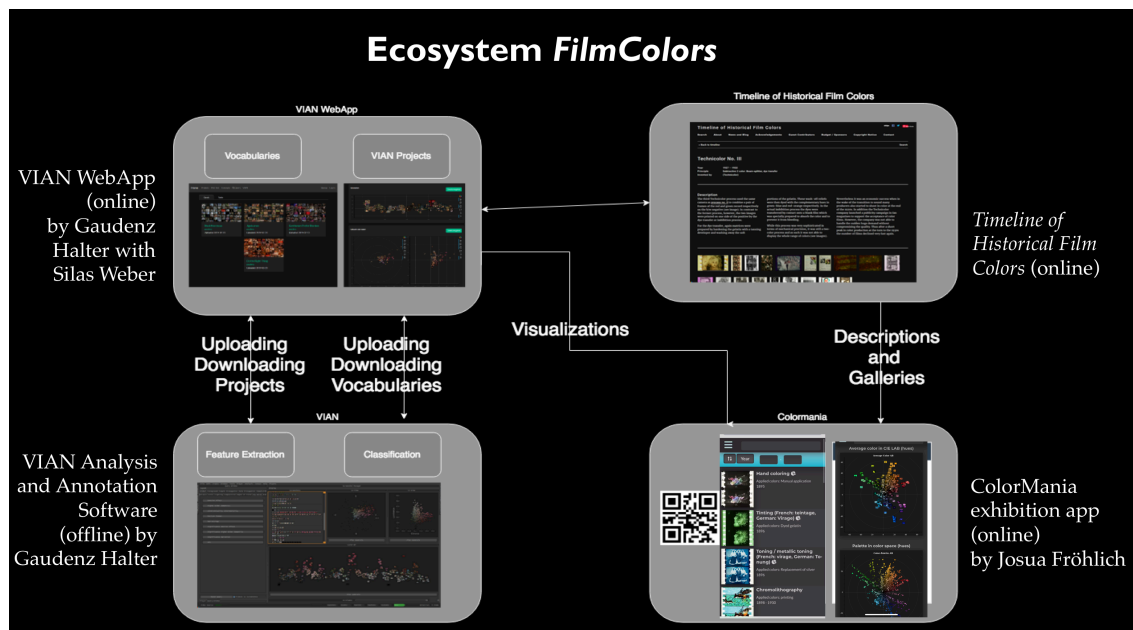


Figure 5. Increasingly VIAN became part of an ecosystem consisting of the offline tool VIAN analysis and annotation software, the online platform VIAN WebApp, connected to the *Timeline of Historical Film Colors* and to the *ColorMania* exhibition app.

All the FileMaker DBs have been exported to the VIAN WebApp database architecture. Each of the individual DBs is thus mirrored in the WebApp.

The corpus DB contains the results listed for each individual film, i.e. on the meso level. It corresponds to the project page on the VIAN WebApp. For each field the corpus DB lists all the occurrences within this film, including a list of all the comments made in the remark fields with the corresponding segment ID. This overview provides an instant footprint for each film and is the basis for hypotheses that lead to further investigations.

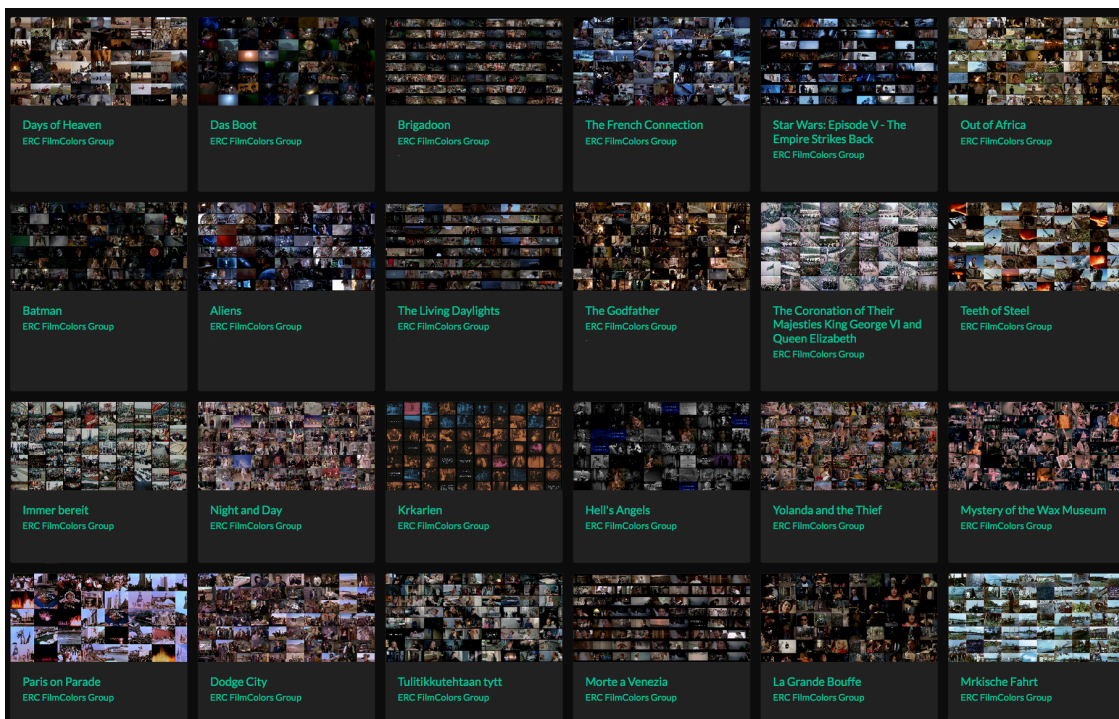


Figure 6. VIAN WebApp project overview of the more than 400 analyzed films. Each project has its own page with detailed analysis and visualizations, see screen video of Paris, Texas <https://vimeo.com/396548709>.

From a different angle, the results are reflected in the glossary DB based on each individual concept, with many custom-made filters for periods, corpus assignment, genres, country etc. With this perspective, it is possible to get instant information about the dominance of a narrative, semantic or aesthetic feature, a motive or location on the macro level, sorted by frequency. The glossary DB corresponds to the concepts page on the VIAN WebApp.

33

Finally the evaluation DB itself where all these results are stored allows the diachronic analysis also on the macro level across the whole corpus. It delivers diagrams of developments over time for the whole period spanning the first 100 years of film history from 1895 to 1995. For instance the development of certain lighting styles such as colored light, mood lighting or mixed lighting, of types of depth of field, of layered and complex image compositions become visible at a glance. These trends are then the foundation for hypotheses that have to be investigated in detail in the analysis DB. It is also crucial to keep in mind — as we will discuss in more detail in the problems section (see Section 3.1) — that these results are not necessarily hard facts. They provide insights into tendencies that have then to be investigated in more detail and tested with other means of evaluation. But the results by far exceed previous traditional, often anecdotal approaches that yielded much less evidence of historical developments.

34

In line with the goal of the research project to identify correlations between technical innovation and aesthetics, the identification of diachronic patterns through digital humanities tools is the single most important foundation for new insights that surpass previous findings. It displays important connections between the technology and stylistic forms but makes also very clear that often such causal connections were overstated in the past while cultural or inter-medial influences were largely neglected. As several recent studies have shown, careful integration of colorimetric visualizations into reflections on the material aesthetics of films yield new insights into the material aesthetics of film, for instance in a recent study of Len Lye's experimental color films of the 1930s [Flueckiger 2019]Flueckiger 2019).

35

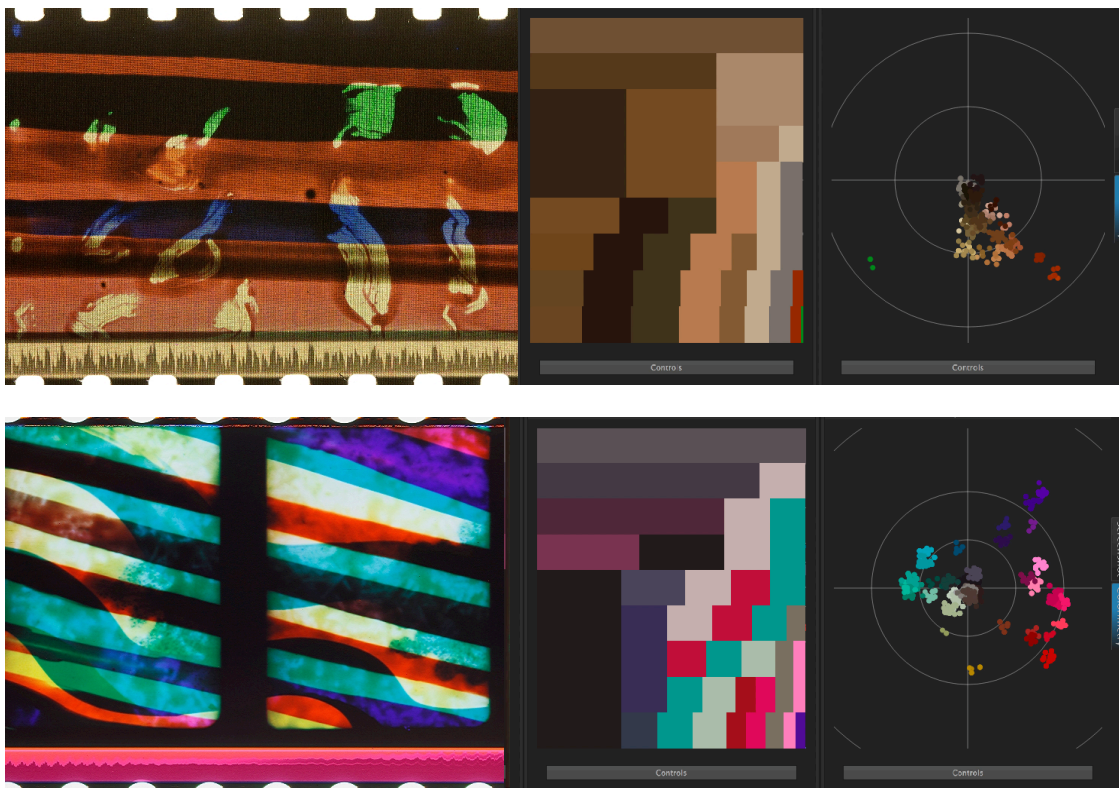


Figure 7. Comparison of colorimetric analysis performed in VIAN gives insights into the material aesthetics of historical color films: Dufaycolor's muted, brownish hues in *A Colour Box* (GBR 1935, Len Lye) vs. high saturation and color separation in a Gasparcolor print of *Colour Flight* (GBR 1937, Len Lye), see galleries on the *Timeline of Historical Film Colors* https://filmcolors.org/filter/?_sft_ubercategory=lye&post_types=gallery.

Increasingly the screenshots themselves became an important part of the investigation. Initially only three exemplary screenshots were embedded into the first version of the glossary DB to illustrate the concepts. Once it turned out that the glossary database is also a perfect way to organize screenshots its architecture had to be completely refurbished to embed the screenshots dynamically via a second database, glossary images DB, all of which was connected to the corpus DB. With this database architecture it became possible to write scripts for portals and to sort images according to periods, corpus assignments or typology. For the VIAN WebApp, it is mandatory to select a sample of the most representative images for external users, which is done by assigning priorities to the screenshots to sort them out. Now the FileMaker DB architecture also allows to embed the screenshots into the corpus DB to provide a selection of the most significant forms of expression through color in a specific film.

36

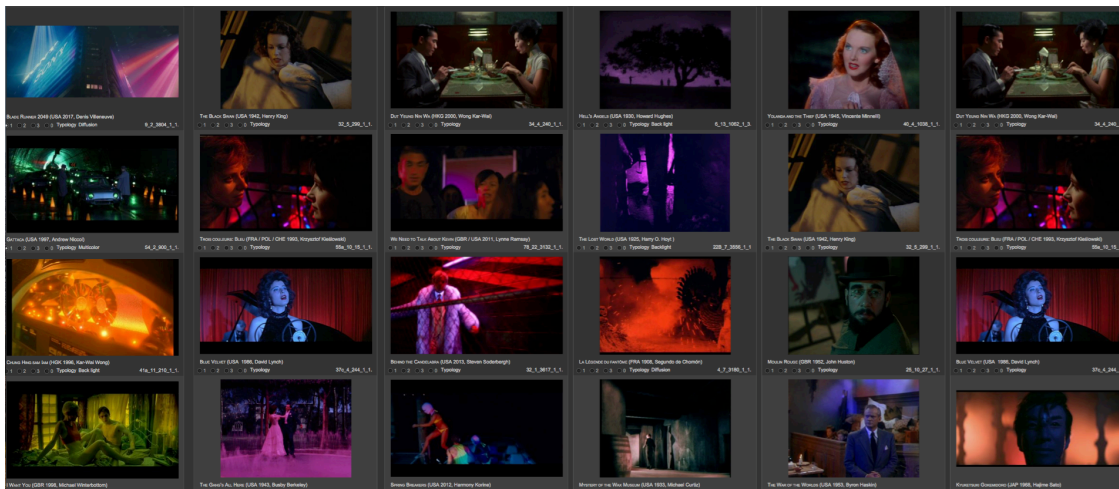


Figure 8. Glossary DB: Extract of a representative selection for the concept "colored light," sorted according to different periods, sub-corpora or typologies.

Since the team aims at capturing historical film prints of the analyzed films to get a better reference for the analyses, these

37

photographs are then published on the *Timeline of Historical Film*. As shown in the VIAN Ecosystem the *Timeline* has been integrated into the workflow as well. A script in FileMaker connects the corresponding galleries from the *Timeline* and displays them in a browser window directly in the databases to compare the photo documentation of one or more historical film prints with the analyzed digitization from DVDs and Blu-rays. The browser window enables immediate frontend tagging of the *Timeline* photos with the thesaurus that is organizing the historical color film processes, the media, quotes, and galleries by a tagging system in the *Timeline* (see Figure 9). Furthermore, the links to the galleries on the *Timeline* are embedded into the project page of the VIAN WebApp.

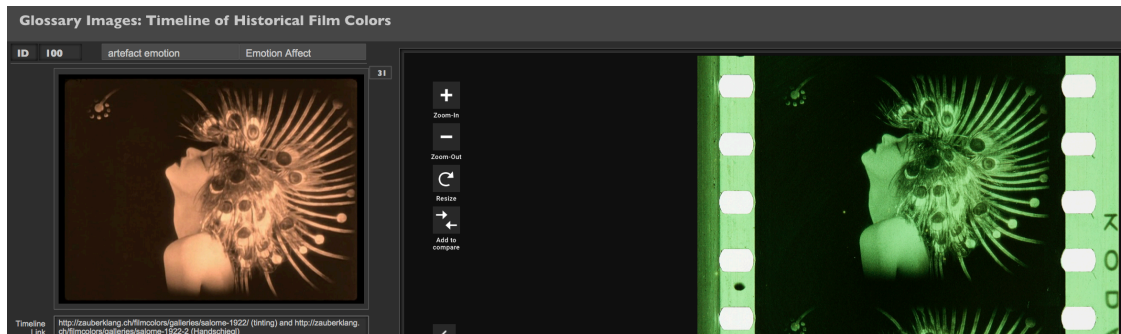


Figure 9. Glossary images DB: Comparison of DVD screenshot (left) with photograph of historical film print of *Salomé* (USA 1922, Charles Bryant) from the *Timeline of Historical Film Colors*, integrated into the DB. Credit: George Eastman Museum. Photograph by Barbara Flueckiger.

Increasingly all the facets of analysis have been integrated into the overarching eco system (see Figure 5 and Sections 4 to 9) that has guided the development of the crowd-sourcing platform and the implementation of all the facets of our research.

3.1 Problems

As mentioned above, resulting from this work was a massive quantity of data and screenshots assembled in a master analysis DB supposed to be hosted on the FileMaker server of UZH. It turned out, however, that the server was not configured for such a demanding task, which necessitated that all the screenshots were exported, down-sized and reimported what seemed like an unsurmountable task for FileMaker due to the non-standardized nomenclature for the image files. Therefore, we ultimately decided to export each screenshot into an individual folder with a defined nomenclature consisting of item ID, image ID and shot ID that were then processed externally by Gaudenz Halter and reimported automatically with a script in FileMaker. Processing included the resizing and compression of the images as well as finding the timestamps of the screenshots within the movie. Since the exported screenshots did not exactly match the content of their corresponding frame, due to resizing and compression earlier in the pipeline, the best matching frame has been determined by application of the mean squared error. This mechanism also allows to import already existing screenshots into a VIAN project and assign them to the correct locus in the video.

Similar problems occurred with the evaluation of the data. Scripts in FileMaker to assemble the data in summations for each film became too complex and the process incredibly slow and vulnerable. Even the in-house FileMaker specialists and external experts could not offer solutions. Therefore, we turned to a similar workflow to export all the data, process them externally in several Python scripts organized in a pipeline. In a first step we had to migrate the complete dataset exported from FileMaker into the VIAN WebApp database architecture. We then calculated the frequencies of keywords on a per-movie basis and related correlation matrices between keywords. This step was followed by a set of successively performed steps to enrich the existing dataset with numeric color features, including the computation of color histograms, color palettes and average color values for each segment and screenshot in a figure/ground separated manner. Finally the per-movie frequencies of keywords have been reimported into FileMaker for the evaluation. Again due to the fact that each team member received their own analysis layer organized with respect to their preferences and interests there were many inconsistencies that affected minutiae such as spelling, local and global concepts etc. Even complicated by the fact that the glossary was extended over time there were internal inconsistencies affecting the connection between the tri-partite logic of the taxonomy in the glossary consisting of classes, fields and concepts, and the organization of the values in fields of the database.

3.2 Lessons learned

To gather consistent and significant data it is mandatory to coach users as much as possible and to illustrate concepts with precise visualizations from screenshots. The glossary database thus contains a priority field to select the most informative and clear-cut

screenshots for each concept, ideally at least six screenshots from different periods for each one of them. As stated before, these screenshots are instantly available on the concepts page of the VIAN WebApp (see Figure 10), so that users get a very good idea what each keyword is referring to. These catalogues of screenshots might be extended by short video clips taken from the corpus in case where movement or other changes over time are central to the concept.

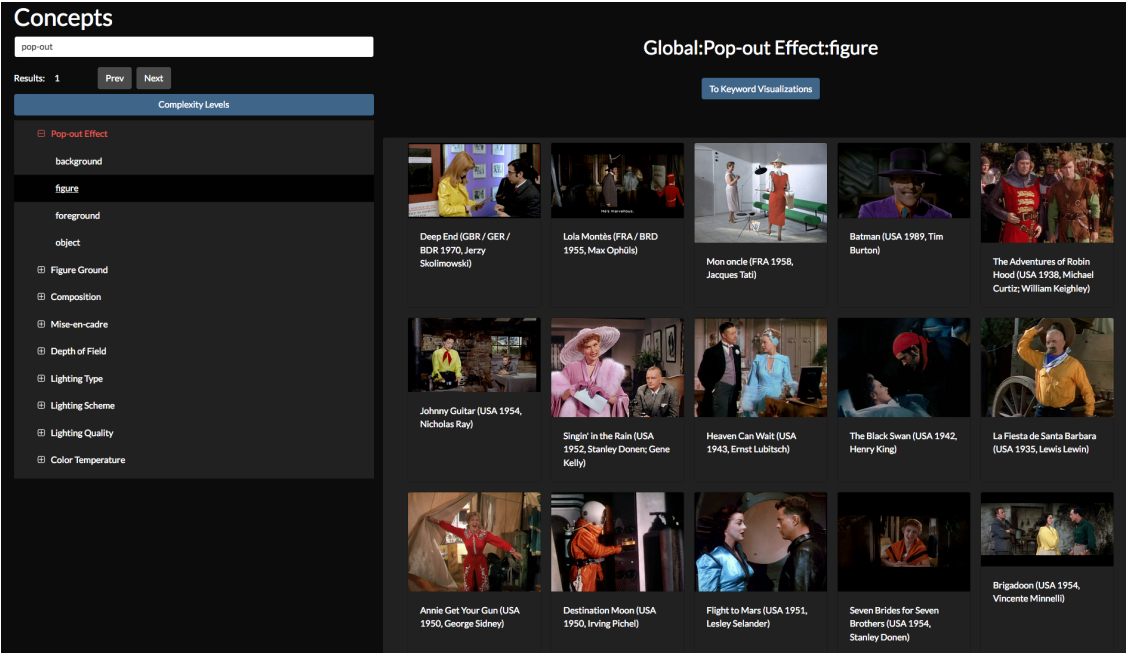


Figure 10. VIAN WebApp concept page: collection of exemplary screenshots for the concept “pop-out effect” attributed to a figure (as opposed to objects).

4 Development of a Visual Annotation, Analysis and Visualization Platform

Based on the manual annotations executed in 2016 and 2017 with the combination of ELAN and the FM DBs, a set of tools for semi-automatic and automatic color analyses and visualization of results have been developed since 2017. These tools make use of computer vision and deep learning to provide meaningful results [Flueckiger 2017] [Flueckiger et al. 2017] [Flueckiger and Halter 2018] [Halter et al. 2019].

Video annotation tools were among the first approaches to apply digital methods to segment and annotate films with a set of tools, see for instance Gruber et al. (2009), investigations in Giunti (2010; 2014), a detailed assessment by Melgar et al. (2017).

In 2016 we executed an extended research into all the available tools, many of which were not running on newer operating systems anymore, due to the termination of funding (see Flueckiger 2017). Finally we decided to use ELAN (see Figure 11), a video annotation tool that offers a great variety of options and is very sophisticated, but was developed with a focus on the analysis of language by the Max Planck Institute for Psycholinguistics in Nijmegen.

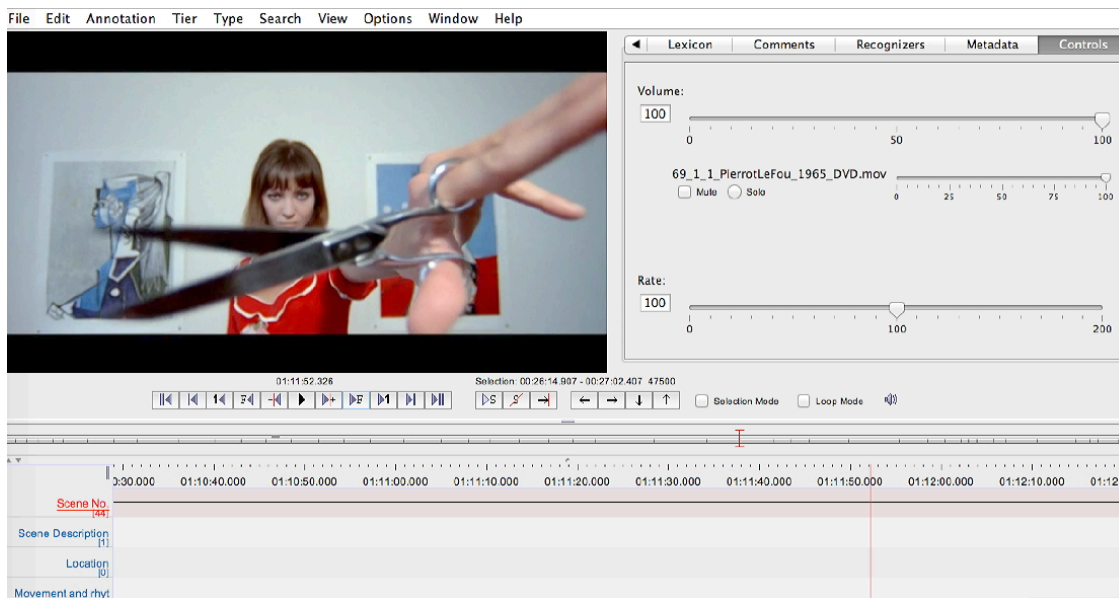


Figure 11. Video annotation system ELAN interface and template for segmentation and annotation, *Pierrot le fou* (FRA 1965, Jean-Luc Godard).

To overcome the limitations of this approach and to shift focus more to the perspective of visual forms of expression, a new visual video annotation system VIAN has been developed by Gaudenz Halter. In addition to several layers of video segmentation and annotation it integrates advanced methods for the analysis and visualization of film colors and is suited for large scale classification of film content.

45

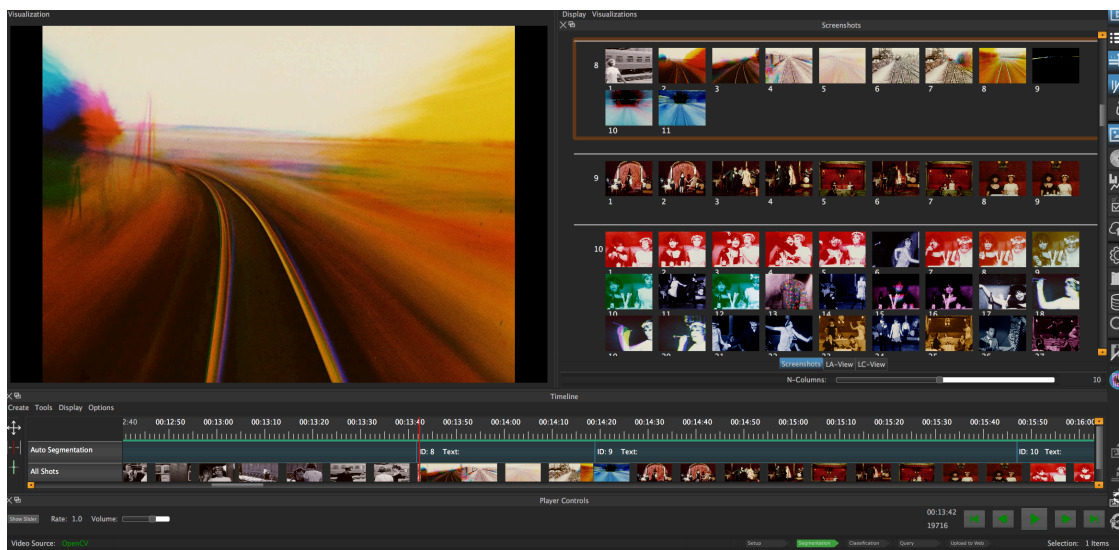


Figure 12. VIAN segmentation layer with screenshot manager, *Sedmírásky [Daisies]* (CZE 1966, Vera Chytilová).

VIAN is a tier-based film annotation software that places emphasis on visual aspects of film style and its color aesthetics, allowing the user to perform general annotation tasks as well as numeric analysis of film material. VIAN has been developed to not only provide data for the crowd-sourcing tool VIAN WebApp that combines our developed analysis pipeline into one software, but also to be flexibly used in other research projects with different film-analytical topics. In essence it consists of several crucial ingredients: Screenshot management, classification by large vocabularies, a toolset for color analysis and visualizations, for a basic overview see several tutorials for the VIAN annotation tool: [vimeo.com/user/70756694/folder/1220854](https://vimeo.com/user70756694/folder/1220854).

46

Previous annotation tools do, to the best of our knowledge, not implement a “screenshot management system”, thus screenshots usually have to be exported and managed by the user in the file system, an obviously difficult and error prone task with an increasing number of screenshots. However, as stated earlier, screenshots play a key role in the visual assessment of films. Therefore, screenshots have become a central type of annotation that can be created in VIAN. Apart from screenshots, VIAN also

47

provides temporal segments and vector graphic annotations. The latter describe annotations that can be drawn directly on screen, currently these are ellipses, rectangles, images, text and free-hand drawings.

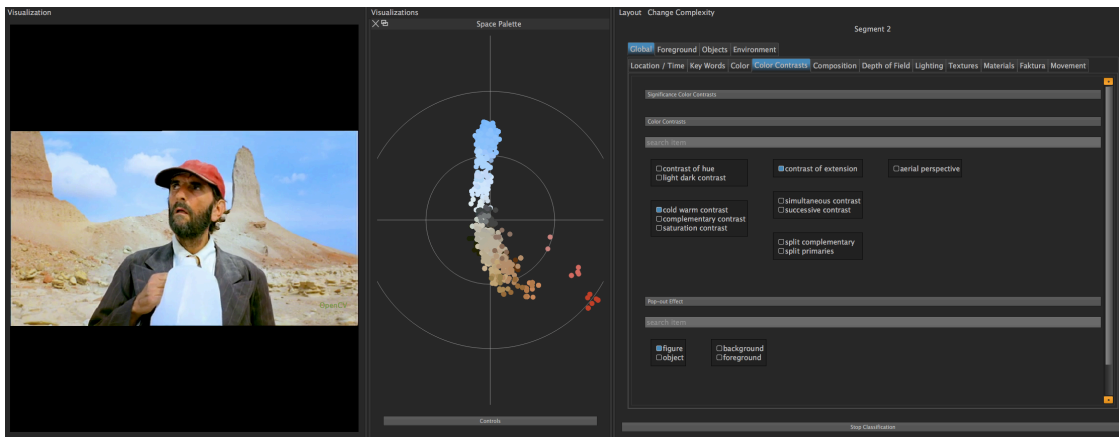


Figure 13. VIAN's analysis widget contains the controlled vocabulary developed during the manual analyses and defined in the glossary DB and on the concept page on the VIAN WebApp. *Paris, Texas* (DEU / FRA 1984, Wim Wenders).

As mentioned before (see Section 2), our project included the classification of a large amount of segments by over 1.200 concepts using FileMaker. With respect to the VIAN WebApp crowd-sourcing tool, this functionality has been implemented into VIAN also. By contrast to many other film annotation software packages, VIAN makes a clear distinction between natural language-based annotation and “classification” based on vocabularies that have been established and tested in the manual corpus analysis. Descriptions are performed by simply typing the respective annotation into the temporal segment or as vector graphic annotation onto the screen. The latter, however, is performed within VIAN's classification system. Once a user has created one or several annotations consisting of screenshots, temporal segments or vector graphic annotations they can be classified by the vocabularies defined in the glossary. VIAN also allows the user to define the conceptual entity, so called “classification objects” that are classified explicitly with one or more vocabularies. For example, the concept “saturated” could target the classification object “male protagonist” and “female protagonist.” Color features can be extracted for an annotation to create visualizations that yield insights into the colorimetric context of screenshots, temporal segments or regions within the frame. Furthermore, VIAN automatically computes several measures in an evenly spaced manner for the complete movie to directly display the most important color features while scrubbing through or watching a video.

Implemented in Python, we have put strong focus into the extendibility such that scholars can easily extend VIAN's functionality to suit the needs of their research questions or using it as a Python API.

4.1 Problems

The development of VIAN has been an iterative process of development and testing with the research team. Obviously, a large number of design questions arise during such a process, especially when the number of requirements and tasks are as large as in the case of the film colors research. Clearly, there are numerous questions related to the software architecture and implementation of tools, but we have also found that developing an easy to use software can be challenging. One of the major difficulties regarding the architecture of VIAN has been to develop a software that solves the very specific need of our research project while remaining generic to be used for other projects and research topics.

Another difficulty was related to the efficient storage of the data. Most annotation tools use a human-readable file format such as XML or JSON to store the generated data permanently. These formats have the advantage that the data can easily be read even without the source tool at hand and improve interoperability. However, numeric data as generated and operated by VIAN had to be stored in a faster accessible file format. During the development we tried several approaches. We started with a simple JSON file. Once the numeric data became too large, we migrated to an SQLite database. This approach did however not scale well enough, finally we implemented a hybrid system using a human-readable JSON file for the annotations and project structure and an optimized HDF5 file for numeric data.

4.2 Lessons learned

We have found that the most important part about the development is a short feedback loop between the developer and the users,

film scholars, students or other researchers. Since there is a huge palette of statistical and analytical methods that could be implemented into VIAN. It turned out that developing in a user-centered fashion is favorable over implementing a large range of possible features. As such we tried to create a solid architectural foundation and remain generic whenever possible without introducing too much complexity.

5 Temporal Segmentation, Extraction and Organization of Screenshots

Approaches to the parsing of films vary greatly depending on a researcher's interest: 53

They can be parsed meaningfully into a hierarchy that has units within units within units [...]. Within this hierarchy, some units have the psychological stature of being events. That is, viewers judge them to have beginnings, middles, and ends, with boundaries that are often denoted by changes in time and place, and that form separable segments within the ongoing audiovisual stream. [Cutting et al. 2012, 1]

What sounds relatively simple, turns out to be rather complex, especially when we consider tools for (semi-)automatic segmentation (Hahn 2009). Ambiguities increase when we define temporal units by the consistency of color schemes, which is the aim of the film colors study. Even if the camera angle varies or if the camera moves in tracking or crane shots the colors dominating the scene can vary significantly, albeit continuously. Therefore it becomes difficult to identify the temporal segments in a consistent way. Some montage patterns — such as parallel montage — require sub-segmentations that consider both event boundaries and temporal units conflicting with each other. While silent films with their intertitles and / or uniformly tinted segments often signpost their structural organization in a rather distinct way, more recent films have more fluidly overlapping scenes. The classical Hollywood continuity system, on the other hand, has established a number of enunciation marks that communicate scene changes or temporal ellipses such as dissolves, fades, or wipes.

Temporal segmentation of a film by human observers — and especially those trained as film scholars or advanced students in film studies — take all these various, historically established cues into account, even if the task is connected to mainly the dimension of film color. On average the team identified between 50 to 70 temporal units with sufficiently consistent color schemes within feature-length films. 54

To accelerate this time-consuming process, VIAN provides an auto-segmentation functionality that computes a temporal segmentation by means of agglomerative clustering of evenly spaced color histograms. The result can then be fine-tuned by the user using the merge and cut tool of VIAN's timeline. 55

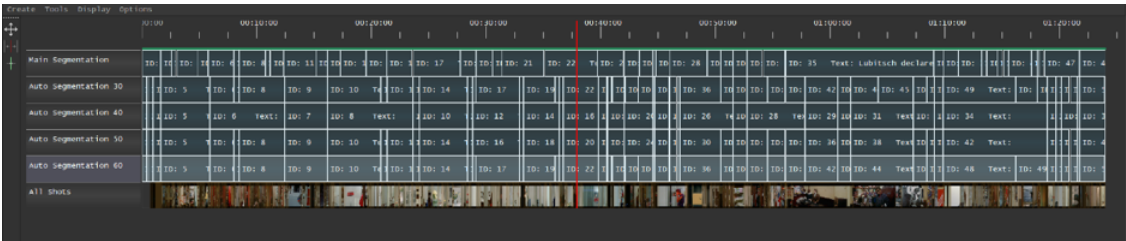


Figure 14. Comparison manual (top) vs. four types of automated temporal segmentation with 30 to 60 segments in *Une femme est une femme* (FRA 1961, Jean-Luc Godard).

As elaborated above (see Section 4), an informed selection of screenshots is paramount as a heuristic tool to reduce the complexity of the time-based video stream by picking out the most relevant moments. Therefore, it became mandatory to establish a fast and flexible process not only to extract screenshots with a simple command but also to organize them instantly in relation to the temporal segmentation of the film in individual bins and with consistent nomenclature. In ELAN, each screenshot extraction required several steps from 5 to 12 commands plus manually naming the image files and defining the image format. VIAN, by contrast, treats screenshots as an integral type of annotation, their creation and management are therefore key functionalities. Screenshots are created with a simple hotkey and displayed in several ways within VIAN including temporal alignment in the timeline and grouped by segmentation in the screenshot manager. 56

5.1 Problems

Many video annotation software solutions have been established in the past that fulfil basic needs. But there is a big leap to developing more sophisticated tools that respond to more complex requirements. Bottom line: the devil is in the details. A very fine framework that integrates several types of players for different zoom-out functions is a powerful start to segment movies 57

temporally, to verbally annotate these units and to extract screenshots. How well does the integrated player process diverse codecs and aspect ratios? What options does it offer to adjust segmentations and to add sub-segmentations for discontinuous entities such as parallel montage? How does it prevent overlapping segments or, by contrast, enable them? What options are there to correct existing segmentations?

Automatic temporal segmentation proved to deliver surprisingly good results that in certain instances challenged human approaches to subdivide the video stream into consistent chunks. On the negative side, auto-segmentation seemed to be much more finely grained in dark scenes and some segments were too long, especially when compared to the average lengths of segments.

5.2 Lessons learned

To develop a robust video annotation system constant user feedback from experienced users is a necessary requirement. For the next step of auto-segmentation we envision to take music cues and sound design into account, see for instance Burghardt et al. (2016) for a very original approach to combine image and dialogue in film analysis with digital tools. Very often onset or termination of diegetic music indicate shifts in locale or time. Sound design is expressive of certain locations or temporal cues as well.

Furthermore, so-called enunciation marks such as fades to black or white, cross-fadings or intertitles should be incorporated into the system of rules for the parsing of units. Significant deviations in the resulting length of segments compared to the average should force the system to process these extremely long or short chunks again.

6 Figure-Ground Separation

Very early in the project, a figure-ground separation tool was established [Flueckiger 2017] to extract characters from the background using a current, deep-learning semantic segmentation technique [Long et al. 2015] [Zhao et al. 2016]. With the rationale to assign to each frame pixel a label this approach indicates the most probable object it represents. It aims at investigating the aesthetics of color attribution through costume and set design in conjunction with other parameters of the mise-en-scène. In the project the method has been constantly improved for speed and performance and provides the basis for all the other color analysis tools — *LAB plots* (see Figures 18 and 21 etc.), *Color_dT* plots (see Figures 15 and 17) — that consider characters independent from their backgrounds.

Aesthetics of figure-ground separation varies greatly during the course of film history, depending on many factors such as color processes, cinematography, mise-en-scène including lighting, staging, materiality of costumes, objects and environments, but also notions of taste and professional norms. For instance strong figure-ground separation was a typical stylistic means to enhance instant legibility in the context of the so-called “continuity system” established in classical Hollywood films from the mid-1920s onward.

In this production context there was often a hierarchy that attributed the most visually compelling colors to the female star and to reduce color difference between characters and backgrounds for supporting characters. Saturation is mostly attributed to female characters while male characters only wore colorful dresses when they were playing certain parts that were framed within cultural norms, by historical distance — for instance royalty or uniforms —, by certain milieus such as the entertainment industry or the arts, by cultural othering such as exoticism or individual personality traits such as queerness or non-conformist attitudes (see Bohn 2000, Vänskä 2017) or genre conventions. Strong figure-ground separation as a trend can be observed again in the emerging contexts of the first auteur-centered color films in European and for instance Japanese productions that feature a sober modernist style.

To investigate these stylistic and culturally justified changes in a clear-cut way we established a typology that took the following dimensions into account: strong vs. weak, silhouettes, figure-ground inversion, and separation by hue, saturation or lightness. By figure-ground inversion we denote relationships where the background is either more saturated or brighter than the background.

These distinctions have then become the underlying concepts for the visualizations that came out of the figure-ground separation pipeline.

Referring to the annotation and classification system explained earlier, VIAN allows the user to define *classification objects* to express a conceptual entity of his or her interest, in this case “figure” and “ground.” VIAN uses a deep learning based semantic segmentation to interpret the content of a frame (see Figure 15). The output of such a segmentation is a grayscale image, where each pixel of the input image is assigned to a gray value, so called *labels*, which correspond to defined objects the model has been trained on. VIAN now allows to assign a set of labels to each classification object, creating a semantic link between the content of

the frame and the classification performed by the researcher. Arnold and Tilton’s studies of visual culture also applied image recognition with deep learning tools [Arnold and Tilton 2020a] [Arnold and Tilton 2020b].

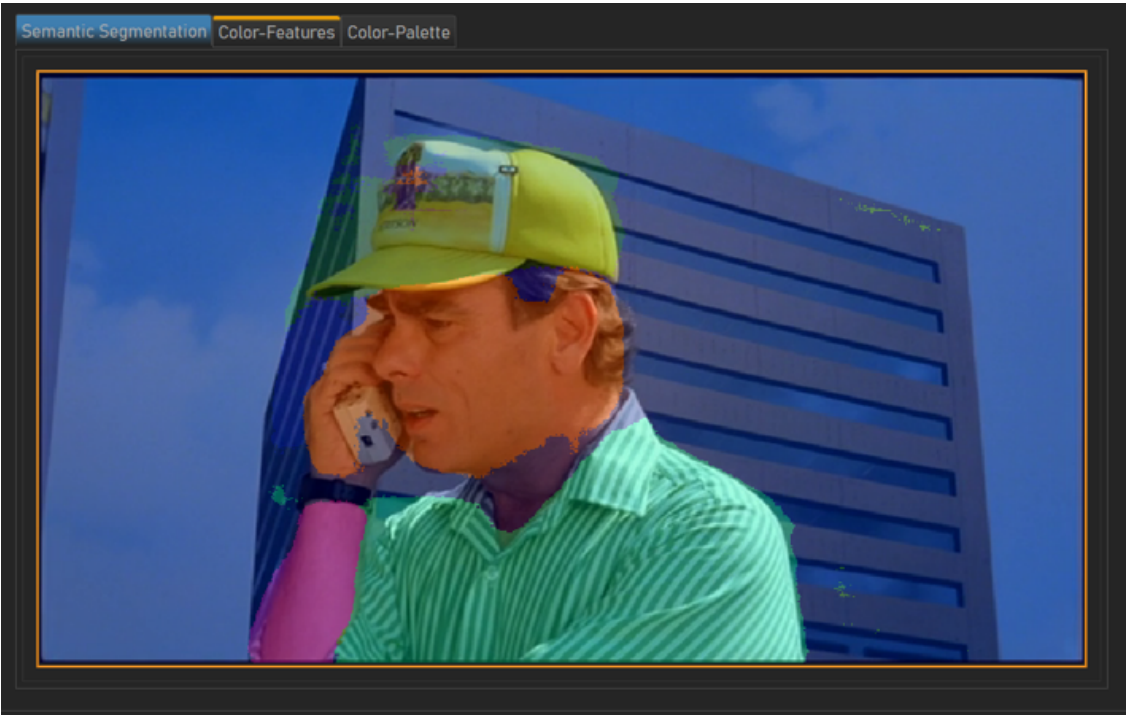


Figure 15. Semantic segmentation performed in VIAN.

The results of this figure–ground pipeline are highly significant, especially when combined with the *Color_dT* visualization (see Figure 15). An instant fingerprint of a film’s aesthetic development emerges when we compare the varying relationships between color attribution to characters vs. environment in the course of a film’s narrative unfolding. As will be elaborated in Section 7 *Colorimetric Analyses and Visualizations*, the resulting types of visualizations differ profoundly from established ones. Mapping the results still raises some questions for scaling. For instance, we found that humans perceive saturation levels attributed to characters as higher when the rest of the image is less saturated, a difference that cannot be rendered accurately with our current visualization and colorimetry methods yet.

67

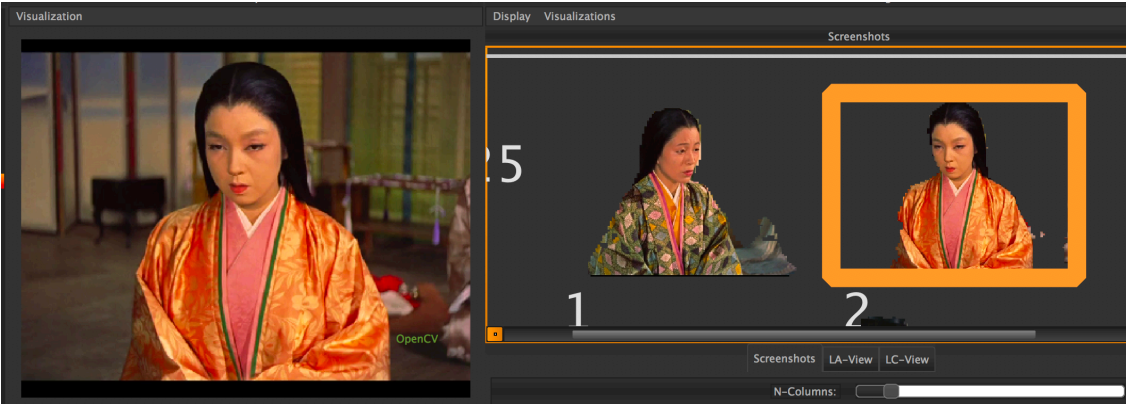


Figure 16. Figure–ground separation in VIAN, *Jigokumon* (JAP 1953, Teinosuke Kinugasa).

6.1 Problems

While we expected this task to be very demanding it turned out that — because this is one of the most important tasks in autonomous driving — deep-learning methods are currently in a very dynamic state especially with regard to extracting characters from backgrounds. YOLO [Redmon et al. 2013] was the first object recognition software applied. It provided very reliable results for the identification of humans while other objects were often misinterpreted, especially when they were partially occluded or cut off at

68

the frame's edges.

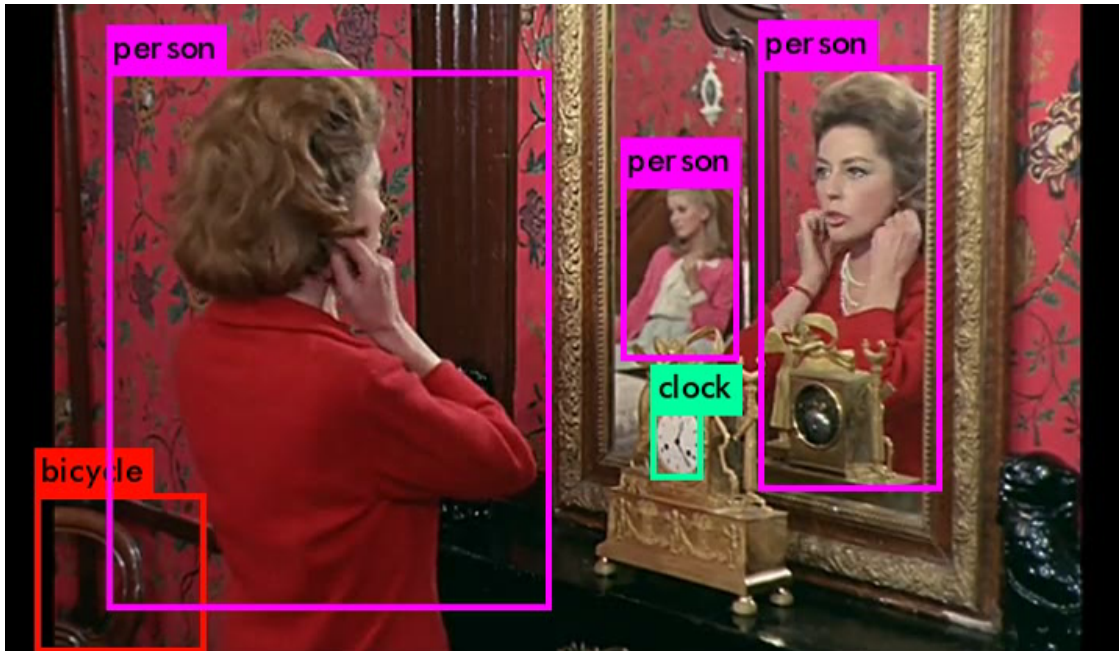


Figure 17. Object identification in YOLO: while persons are identified consistently, errors occur when objects are partially occluded or fragmented.

YOLO was combined with GrabCut [Rother et al. 2004]. GrabCut works as follows: the user initially draws a rectangle around the object to be in the foreground, GrabCut will then try to directly segment the frame, and return the result. The user can then iteratively optimize the result by marking regions that have not been identified correctly using strokes. Performing this process for each image manually would not have been feasible because of the time constraints, we thus used YOLO, an object recognition neural network to draw the initial bounding box and the strokes. However, this pipeline did not scale well enough for our purposes, demanding a large amount of resources and time. We therefore decided to use a deep-learning convolutional network to perform the pixelwise segmentation directly with semantic segmentation [Long et al. 2015] rather than the YOLO and GrabCut based approach.

69

6.2 Lessons learned

A collaborative, interdisciplinary approach that connects high levels of expertise both in the domain of aesthetic analysis and computer science has proven to be mandatory for the elaboration of an analysis and visualization pipeline that respects both fields and connects them in a convincing manner. While such a statement may seem banal, in fact the actual exchange between different disciplines has been much more demanding and requires continuous adjustments from both sides. Experts from the field of humanities must be able to understand the requirements of informatics and to describe the task in a highly formalized manner. Scientists on the other hand need to be open to integrate a sense for the subtleties of aesthetic concepts to understand why minor details unexpectedly can have a significant impact on the results. The resulting pipeline should produce visualizations that respect the rigorous demands of science while also considering instant accessibility for human observers and knowledge of aesthetic distinctions at the same time.

70

7 Colorimetric Analyses and Visualizations

Previous approaches to visualizations of color schemes were surprisingly reduced in their scope and were not sophisticated enough to do justice to aesthetic subtleties of color design in film.

71

Currently available tools to devise color schemes are often applying K-means [Brodbeck 2011] [Rosebrock 2014] and thus are limited to the depiction of a fixed set of hues. Color schemes in VIAN, by contrast, are extracted to match spatial distribution and can be edited according to the needs of the color analysis for a certain film. Some films apply very distinct hues to their color schemes while others resort to minute shifts to display developments in character relationships. Color schemes can express a character's inner states or personal developments, relationship to other characters or a given environment, again norms of taste and milieus, or cultural conventions. Socio-political markers indicate characters' connection to a certain class or social function in a

72

socially or culturally pre-defined way as for instance in uniforms.

From the start, we therefore envisioned a different approach that allows for a flexible fine-tuning of color schemes to match the specific style of a given film. A second basic requirement was the representation of the spatial distribution of colors in a way that is instantly displaying the quantitative allocation of colors in an image or temporal segment. Thirdly, we aimed at visualization methods of color schemes that show their development in a film over time according to the temporal segmentation executed in the pipeline.

73

Typical time-based representations such as movie barcodes or mosaics provide plenoptic overviews of films (for a discussion see Heftberger 2016, Stutz 2016, Olesen 2017, Flueckiger 2017) but they do not represent the finely grained shifts and relationships that are fundamental for an in-depth study of aesthetics. Frederic Brodbeck arranged color schemes in circles to give an overview of what he called a fingerprint of a film's color scheme [Brodbeck 2011]. Z-projections have become a main part of Kevin L. Ferguson's visualizations [Ferguson 2013] [Ferguson 2016] who also proposed a volumetric approach to visualize an entire film's color on the time axis in 3D [Ferguson 2015]. James E. Cutting and his team at Cornell University devised many methods to visualize movies, among others a movie barcode that implemented color schemes from warm to cold colors [Cutting 2016]. From 2013 onwards Lev Manovich [Manovich 2013a] [Manovich 2013b] and his Software Studies lab applied a range of visualizations to Dziga Vertov's films for Adelheid Heftberger's research project [Heftberger 2015] [Heftberger 2016], some of them based on ImagePlot and ImageJ that were used previously for the visualizations of artworks [Manovich 2012] [Reyes-García 2014] [Reyes-García 2017]. ImageJ, initially introduced for bio-medical research [Ross 2007], has since been used by several researchers for film analysis (Olesen 2016, Heftberger 2016, see several chapters in Hoyt et al. 2016). Casey et al. compared temporal segments in films based on histograms visualized in a similarity matrix [Casey and Williams 2014].

74

As elaborated in Halter et al. (2019), the team defined a set of requirements for the visualizations. They should

75

1. Represent visual impressions true to human perception;
2. Represent subtle aesthetic nuances in figure and ground separately;
3. Visualize the films at the micro (screenshot, temporal segment), meso (individual film) and macro (corpus) levels.

And in addition they should be interactive and flexible for adjustment to an individual researcher's interest [Halter et al. 2019, 126].

Therefore, as elaborated in previous papers [Flueckiger 2011] [Flueckiger 2017], the relationships of colors need to be mapped into a perceptually uniform color space to provide visualizations that match human vision. In VIAN, both the screenshots and the color schemes are thus transformed into the perceptually uniform CIE L*a*b* (referred to as LAB in this paper) color space needed for meaningful representation of the color distribution in a given film. Contrary to most established visual representations such as image plots, z-projections, color palettes or barcodes, a visual representation in a perceptually uniform color space pays attention to the relational nature of colors with regard to the visual system. Chromaticity and lightness plots provide an overview of a film's color distribution, see Figure 18.

76

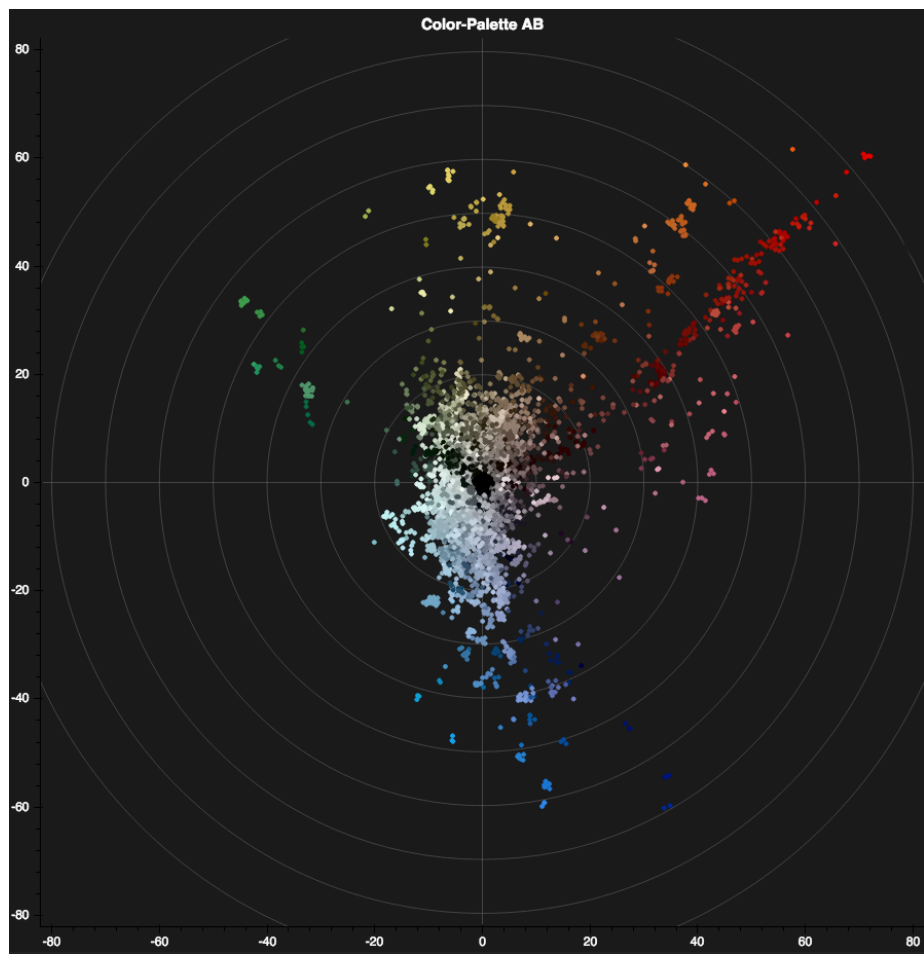
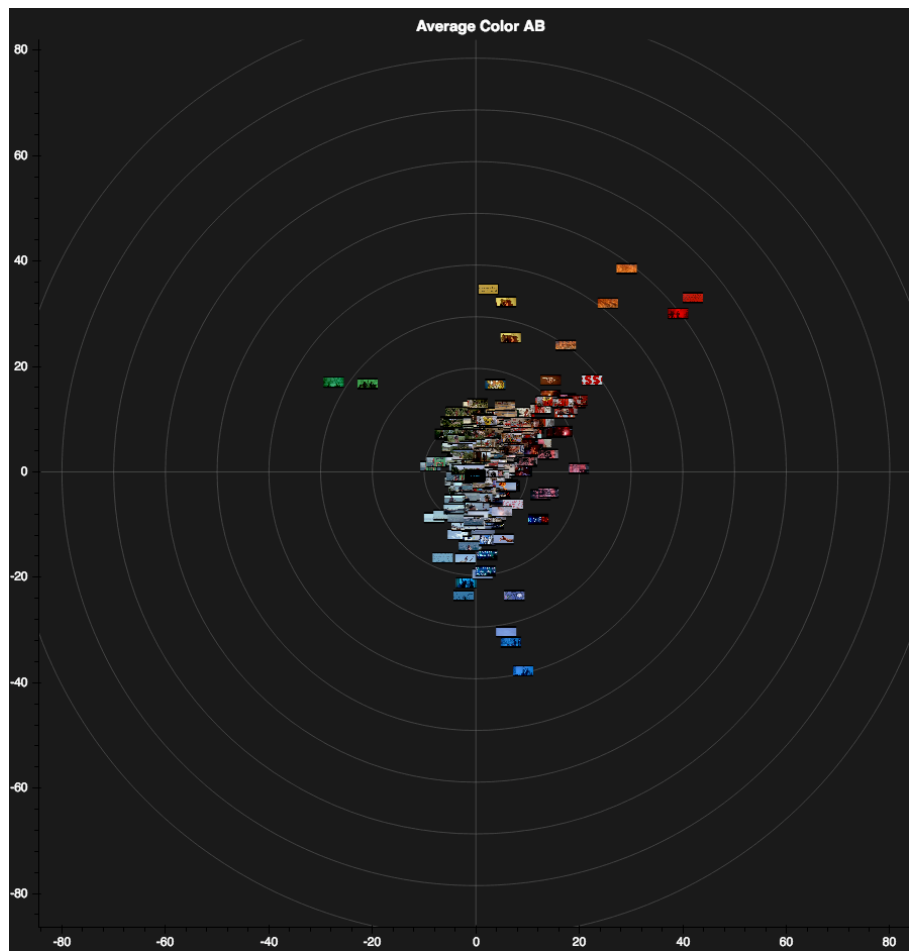


Figure 18. Chromaticity plots in CIE L*a*b* (LAB) for *Pierrot le fou* (FRA 1965, Jean-Luc Godard), image plot (above) vs. palette dot plot (below) The comparison shows that chroma extends much further in palette dot plots without averaging effects caused by the representation of images.

77

While visualizing color schemes in a strip of color patches sorted by frequency is generally well established and gives a good overview, they are often hard to compare and hide how the palette has been assembled during the clustering process. To compare color distribution in relation to human perception the color scheme is displayed in the LAB color space as a palette dot plot. With this method small changes as well as color contrasts within the chroma or hue between different color patches become directly visible (see Figure 19, right). The tree palette (Figure 19, above, middle) should help the user to understand into which final cluster the colors of the input image have been merged. To this end, palettes are stacked in different merge steps corresponding to increasing levels of granularity on top of each other and the color patches sorted within the palette by the order resulting from the clustering. Thus all colors merged into a cluster are visualized directly below it.

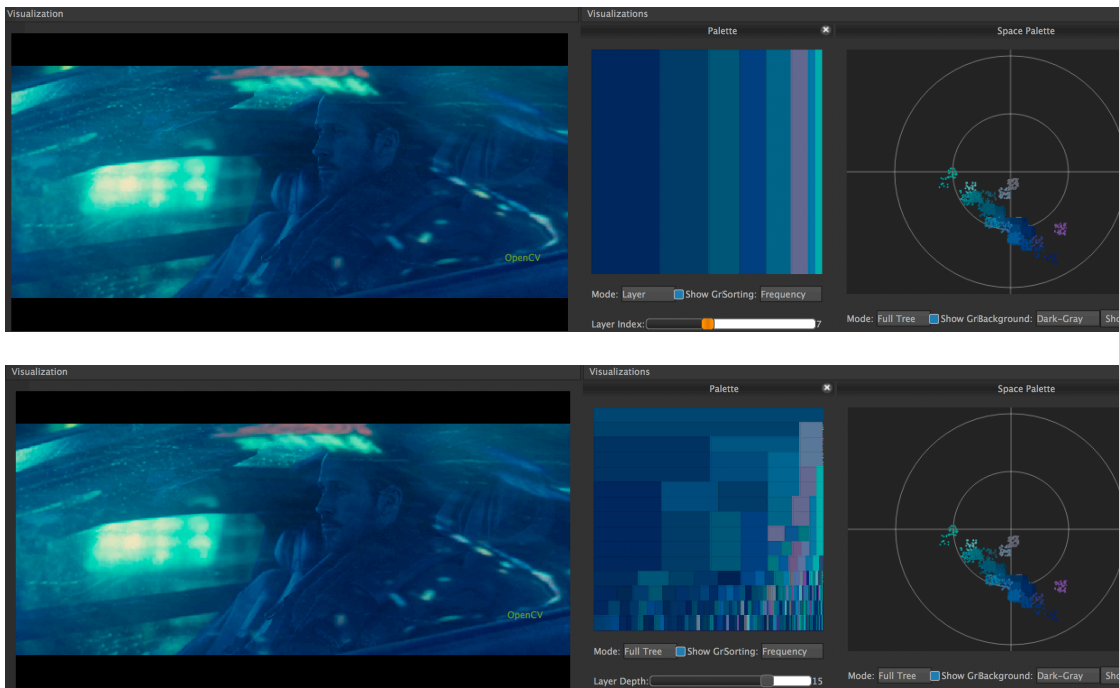


Figure 19. Color palettes for individual shots: tree diagram with increasing levels of detail (above, middle) and LAB (right); selection of 7 hues for layer palette sorted by frequency (below, middle). *Blade Runner 2049* (USA 2017, Denis Villeneuve), see interactive visualization methods in the screen video <https://vimeo.com/299804415>.

78

Applying color-related computational methods, such as clustering or statistics on the raw frames of a film is often not feasible because of the sheer amount of data. It is thus often a necessity to extract feature vectors adequately representing the content through a color histogram that is regularly used within VIAN. However, color spaces are three-dimensional and so are their color histograms, making visualization of color histograms a difficult task. A naive approach would be to visualize the histograms as point clouds in a three-dimensional space but this method doesn't yield good comparability. We therefore developed a bar-chart like representation of a color histogram by sorting the colors of the three-dimensional histogram into a one-dimensional list using a room-filling curve, namely the Hilbert curve. Intuitively, this room-filling curve describes a path, by which any point of a given space is visited, in our case the bins of the three-dimensional color histogram. By unraveling this curve, we can align the bins of the three-dimensional color histogram in a one-dimensional row. We use Hilbert curves, because this type of room-filling curve has shown to preserve the specific locality well, in our case this means that color bins which are close in the three-dimensional histogram, will also be close in the unraveled, one-dimensional, histogram bar plot.

79

Color_dT is an advanced method to visualize the color development of a film over time on the meso level with regard to its temporal unfolding, again for figure, ground and whole screenshots independently. It is currently implemented for saturation contrasts, contrast of hue, chroma or light-dark contrast, but could also include cold-warm contrast. Shifts in figure-ground relationships become instantly evident, so do overall developments with regard to the narrative events in the course of a film.

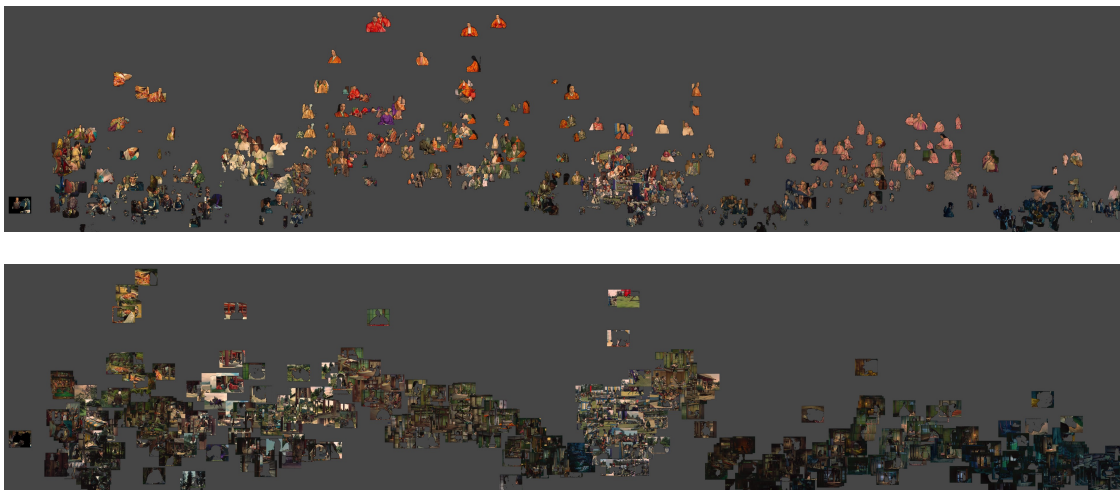


Figure 20. Color_dT plot development of saturation (y-axis) over time (x-axis) for figure (above) and background (below) in *Jigokumon* (JAP 1953, Teinosuke Kinugasa), visualization by Noyan Evirgen for ERC Advanced Grant *FilmColors*.

One significant example is the Japanese film *Jigokumon*, produced as one of the first Japanese color films shot in the then new chromogenic process Eastmancolor. In the first half of the film we see a pronounced figure ground separation with the characters, especially the female love interest standing out in colorfully patterned, saturated kimonos in front of subdued backgrounds. In the middle of the film a peripety occurs during a horse race where the two male opponents fight each other. This scene is set in broad daylight with conflicting colors in background and foreground. After this turning point, the tragedy sets in with a markedly different color design and mise-en-scène characterized by dark scenes in low-key lighting, which by its very nature reduces figure–ground separation.

80

With early applied colors such as tinting and toning, the LAB chromaticity plots look decidedly different due to the mostly monochrome color schemes. As becomes evident from a comparison between *L'Inhumaine* (FRA 1923, Marcel L'Herbier) and *Das Cabinet des Dr. Caligari* (GER 1919, Robert Wiene), the digitization of *L' Inhumaine* differs substantially by the detached distribution of chroma — there is no continuity from the center to the higher levels — which could result from problems in digital color management.

81

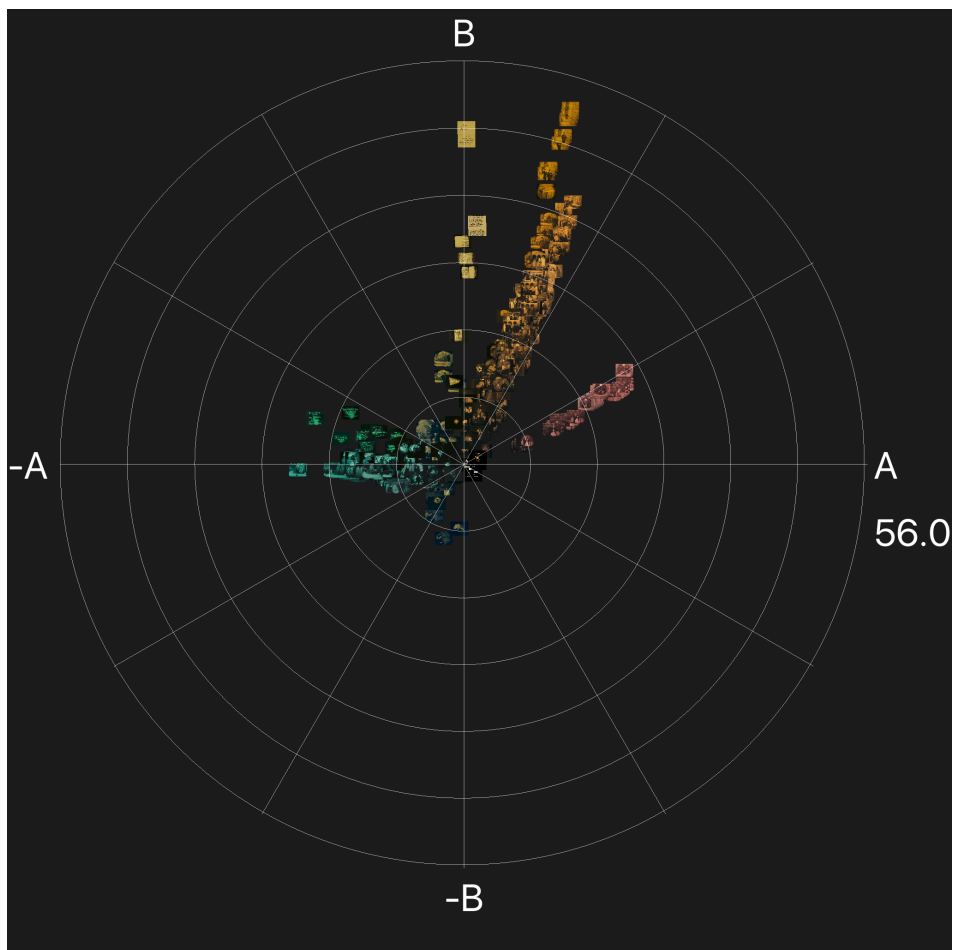
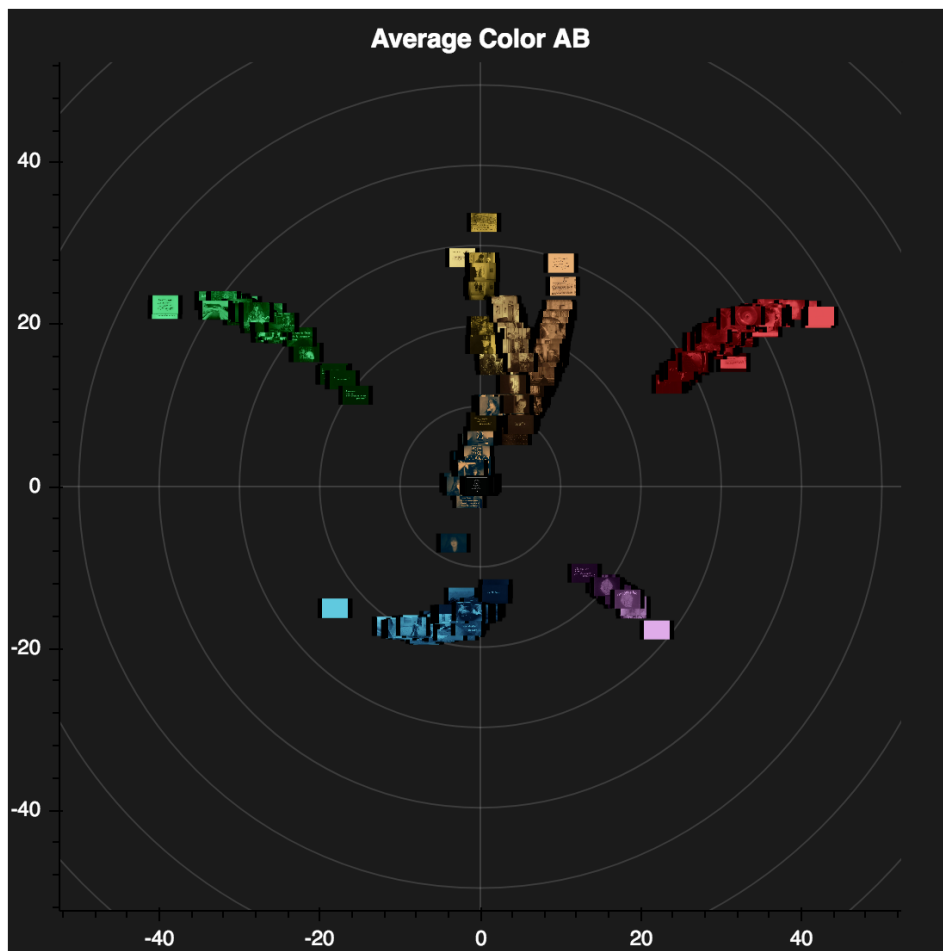


Figure 21. Chromaticity plots in the CIE L*a*b* space for the tinted films *L'Inhumaine* (FRA 1923, Marcel L'Herbier) (above) and *Das Cabinet des Dr. Caligari* (GER 1919, Robert Wiene).

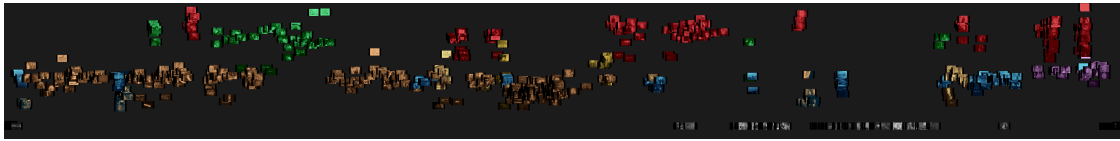


Figure 22. Color_dT plot development of saturation (y-axis) over time (x-axis) for the tinted film *L'Inhumaine* (FRA 1923, Marcel L'Herbier), a tinted film.

Features Tool

Correlations between concepts are displayed in two ways. The *features tool* enables users to select the concepts from the menu. Consequently the occurrence of these concepts are then displayed over time related to the segments where they occur. Connected to the exemplary screenshots this type of visualization instantly builds the foundation to establish and test hypotheses.

82

The correlation between different keywords within a project or corpus-wide can be investigated using the co-occurrence matrix plots, which indicates how often every combination of keywords occurs within the scope.

83

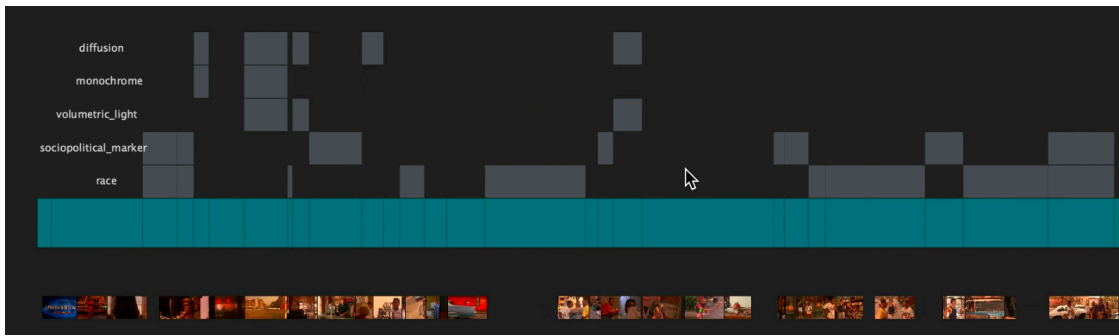


Figure 23. VIAN features tool visualizes co-occurrences of concepts organized on the time axis with regard to the temporal segmentation in *Do the Right Thing* (USA 1989, Spike Lee), see screen video <https://vimeo.com/292861139>.

In the screen video the features tool is tested with Spike Lee's *Do the Right Thing*. Several typical features of this film have been selected in this visualization. In addition to the leitmotif that establishes the dominant red spectrum of the film and associates it to the topic of heat in its double sense as a temperature and as a metaphor for the rising racial tensions, the film's aesthetics is informed by a dichotomy between the private sphere and the public space. The private sphere in interiors is often shown suffused in atmospheric diffusion again associated to the hazy damp caused by the heat in warm monochrome red tones with shafts of light filling the room, all of which are associated to a romantic tradition dating back to the early 19th century. By contrast, the aesthetics of the film's public sphere follows a much more sober style connected to traditions of social realism with extended depth of field in wide-angle shots that show the characters in relationship to each other and to their environment. In terms of color design, the film makes use of what we call socio-political markers, culturally established conventions to denote certain social strata or official functions. For instance the protagonist played by Spike Lee himself is a pizza delivery boy and wears clothes in the colors of the Italian *tricolore*, white, green and red. His encounters are also defined by the central topic of race that is connected to the various ethnic groups, the Puerto Ricans, the Jamaican, the Koreans, the Italians, and the Afro-Americans, each of which is associated to different sets of hues by socio-political markers. Such a pattern of color aesthetics and meaning can easily be confirmed or further elaborated with the *features tool* (Figure 23).

84

7.1 Problems

Image plots are fantastic tools for visualizations when a researcher aims at keeping the connection to the source material. By zooming into the plots, users can look at the screenshots and see where and why a certain screenshot is present in the plot and how it is related to the film. However, because image plots' colorimetric values are calculated based on the average of the screenshot's color distribution, monochrome color schemes distort the visualization by being too dominant. Screenshots with more than one hue or a multitude of hues aggregate at the center of the LAB visualization or at the bottom of a *Color_dT* plot. Therefore, this type of visualization is best suited for early applied colors such as tinting and toning with their monochrome color distribution or

85

for films with stark color designs in mostly one dominant hue per segment such as for instance *Suspiria* or Slawomir Idziak's camerawork whose signature style often applies colored illumination and monochrome or graduated lens filters.

7.2 Lessons learned

For many films that are not rendered well in image plots we devised an alternative solution by consecrating the full image rendition and separated the individual color values (comparison see Figure 18). That is, instead of using the average color values, we computed the color palette for a given screenshot and visualize it in the AB plane of the LAB color space. A jitter effect is applied to add some noise, making the amount of a specific hue visible within the color space. These palette dot plot visualizations now show the color schemes represented by dots for each of the colors present in a screenshot. We had to devise a method to include the spatial percentage into the dot plots. Dot plots have also become a means to show color schemes in a different way than with the typical color bars, see Figure 19. Different methods to scale and distribute colors in visualizations are offered such as zoom functions or range adjustments. Rotation is crucial for visualizations related to the L axis in the LAB space to show the distribution of hues in a meaningful way. While palette dot plots display a film's color distribution in an intuitive way, they do not take the relative incidence into account. Therefore yet another type of visualization was introduced: heat maps that show the color distribution by means of levels of transparency corresponding to the incidence (Figure 24).

86



Figure 24. Palette dot plots (left) vs. heat maps (right): while palette dot plots visualize the occurrence of a certain color, heat maps indicate the incidence by different levels of transparency. VIAN WebApp query “disgust.”

8 Visualizations, Concepts and Correlations on the Macro Level

One of the biggest gains of our investigation is the massive dataset created by the analysis team. As written above, (see Section 3) it amounts to more than 17.000 segments with more than 170.000 screenshots for more than 400 films, each of which are connected to the meticulous manual analysis and annotation presented in detail in the previous sections of this paper (see Sections 2 to 5).

87

One way to display the amount of associations between different keywords within a dataset this large and complex, is to follow a network visualization approach. Every keyword is represented as a node and its connections to other keywords as edges. The more these keywords appear in the same segment the closer they are placed together within the network using the Fruchterman-Reingold force-directed graph drawing algorithm provided by the NetworkX Python library [Hagberg et al. 2008].

88

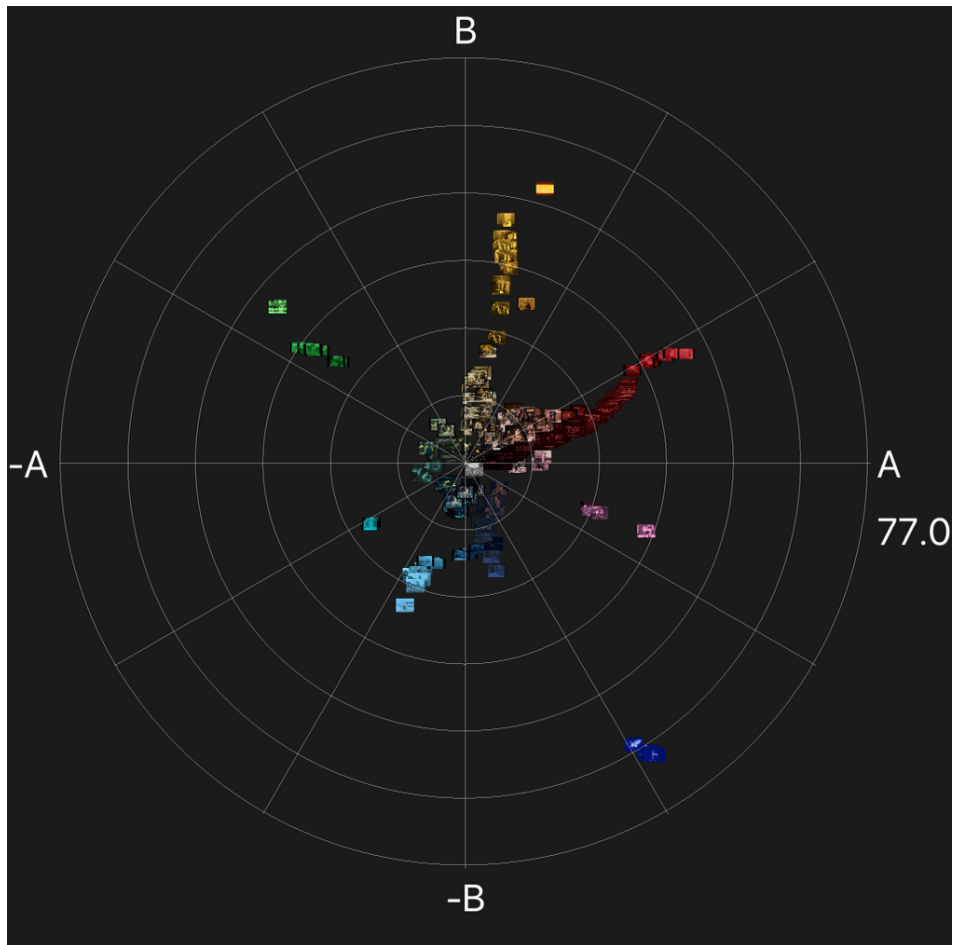
With the integration of this dataset into the VIAN WebApp we open up a broad range of opportunities for queries on the segment, film and corpus level to combine the manual annotation with all the colorimetric analysis and visualization methods elaborated in sections 6 and 7 (see Section 6, see Section 7). By such a comprehensive approach we enable users to combine all the three different levels, from the micro level (close reading, for instance individual screenshots or segments) to the meso level of individual films to the macro levels (distant reading) of the full corpus or selected subcorpora. Such selected corpora can be queried by any

89

concept regarding narrative aspects, characters' emotional states, motives or themes, and all the aesthetic and stylistic dimensions mentioned in section 2 (see Section 2).

Two concept queries are displayed here, the search for dream sequences in the three periods 1895–1930 and the search for night sequences in early film. When we compare the visualization of dream sequences in early film with the period from 1930–1955 two insights emerge, dream sequences are often marked by monochrome color schemes, and often the dominant color is red. In the second plot (1930–1955) the relatively prominent incidence of green is related almost exclusively to the *Wizard of Oz* where the concept of dream applies to the primary narrative of the film (see Figure 25 below).

90



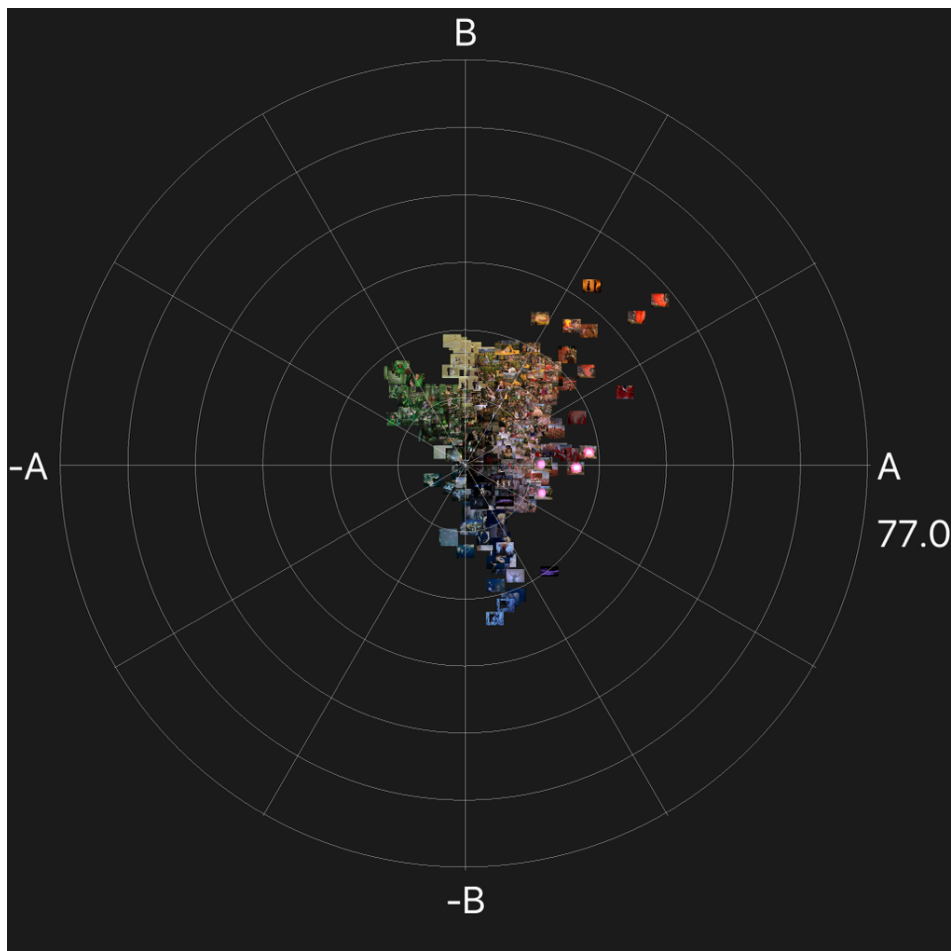


Figure 25. Visualizations on corpus level for the narrative concept “dream,” 1895–1930 (left) and 1930–1955 (right), AB image plots.

Applied colors in films produced during the first three decades of film history — such as tinting and toning with their monochrome color schemes — followed loosely a set of conventions, which then have to be tested in individual films or over a certain period. Because there were many ambiguities, each film’s color schemes and attribution of hues to different locations, times, narrative strands, genre or gender conventions has to be carefully investigated for film scholars and restorers alike to understand the guiding rules of one particular historical film print [Ledig and Ullmann 1988] [Mazzanti 1998] [Mazzanti 2009]. For instance *Das Cabinet des Dr. Caligari* has survived in five differently tinted and toned versions, see gallery on the *Timeline of Historical Film Colors*,^[1] but a reference print of the initial German premiere version has not been found yet #wilkening2014, see Figure 9 for the comparison of a DVD vs. historical print.

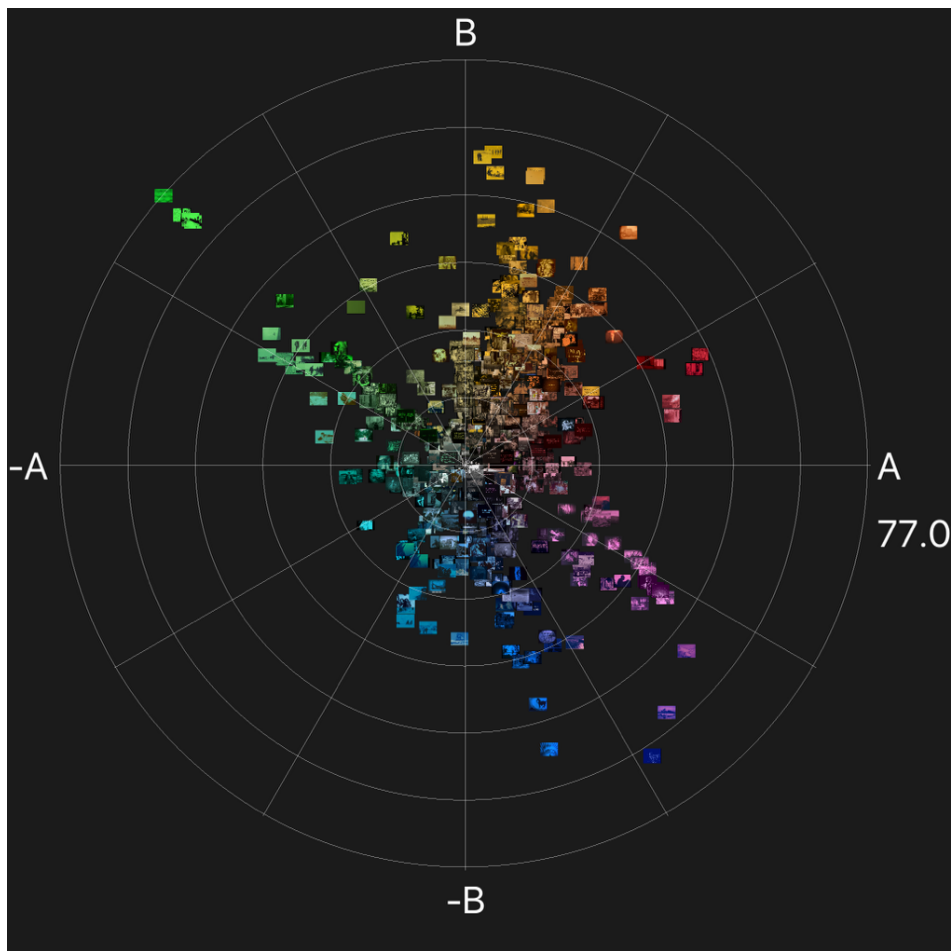


Figure 26. Corpus visualization exterior night in films from 1895–1930, with fire scenes excluded.

One of the most stable associations of specific hues to a certain narrative dimension is blue tinting to exterior night scenes, because limited speed of early film stocks did not allow for night scenes to be actually shot by night. Therefore these scenes needed to be marked by typical hues. The visualizations show that blue is indeed one of the dominant hues with green almost as wide-spread as blue. Amber and red are the third dominant range. Amber is often associated to tungsten or candle light in interior scenes, so segments that contain connections between interior and exterior scenes in a certain sense contaminate the result. Fire scenes, by contrast, are typically tinted in red, so they were eliminated the query, see Figure 20.

In general, LAB image plots and palette dot plots are limited in informative value on corpus level as opposed to their usefulness on the film level, especially when displayed in print. They only indicate trends that then have to be confirmed by looking deeper into the films and segments where they occur. To this end, all the visualizations on the query page of the VIAN WebApp are highly interactive. When hovering over the plots, the researcher gets shown the corresponding segments of the film including screenshots and a scene description. In addition, all the segments and films are displayed with the corresponding color palettes in the form of coarse barcodes.

See screen video of the query page: <https://vimeo.com/402360042>

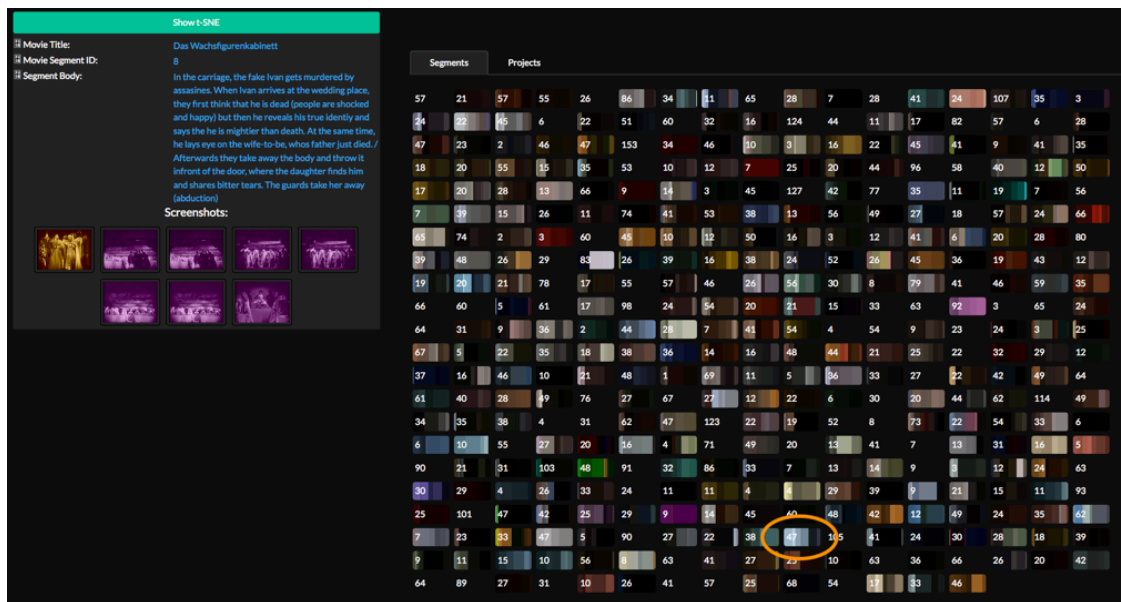


Figure 27. On the query page of the VIAN WebApp, all the segments and projects detected in the query are displayed with a coarse color palette (above). By clicking on a segment, the screenshots, a short scene description and a summary visualization are displayed. *Possession* (FRA / DEU 1981, Andrzej Żuławski).

To investigate the diachronic development, an additional method for corpus visualizations called *Color_dY* was implemented that considers the temporal distribution over years instead of plotting a selected period into an overview in LAB that obscures the color schemes of individual films.

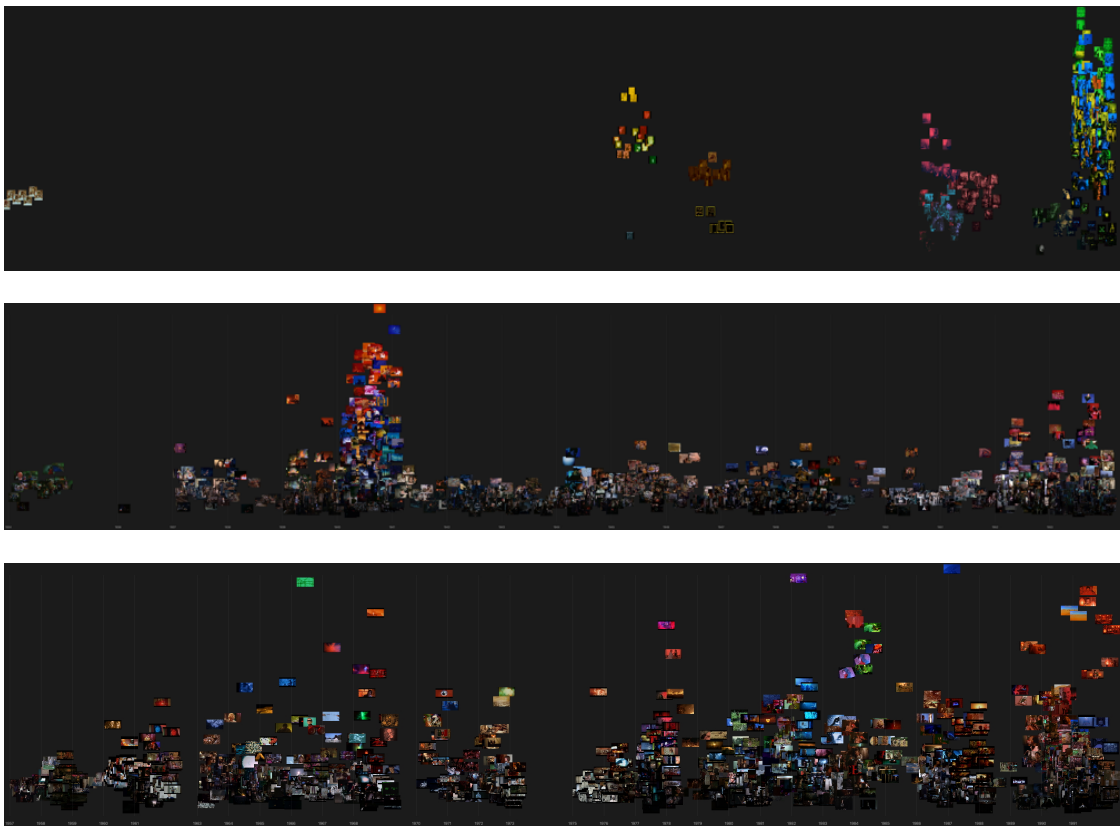


Figure 28. Color_dY plots saturation (y-axis) over time (x-axis) for colored lights, 1895–1930 (top), 1930–1955 (middle), 1955–1995 (bottom).

For instance in the middle plot the animation film *Fantasia* (USA 1940, James Algar et. al.) sticks out with extremely high levels of saturation, and again also for an animation film *Die Abenteuer des Prinzen Achmed* (GER 1925, Lotte Reiniger; Carl Koch) in the first plot on the right hand side. Film titles, segments and screenshots in combination with a scene description are again displayed by a hover function (see Figure 27).

96

8.1 Problems

In the course of developing these visualization methods on corpus level we noticed difficulties to receive clear-cut pictures. One of the problems resulted from the fuzziness of the concepts that generated quite a high amount of noise, as elaborated in the previous section, see Section 7. However the most persistent issue that has been identified is the dominance of monochrome color schemes in the LAB visualizations, in the same fashion as in the image plots per film discussed in the previous section (see Section 7.1). Because of high levels of chromaticity in some monochrome screenshots as compared to averaging effects by variegated hues these images always stick out and therefore distort the result. This effect is even stronger in image plots that represent data on corpus level, because of the variations in different films' color designs.

97

8.2 Lessons learned

One of the first measures we took was to clean up the data. Secondly we also integrated the dot plots explained in the previous section (see Section 7) into the corpus visualizations. One of the most helpful parts of visualizations on corpus levels is the integration of the temporal segments including scene descriptions and screenshots in a sidebar next to the plots. Since the relationship between different keywords becomes complex in such a large corpus, we have implement more types of visualizations to convey the correlation and connections between concepts, the color-features associated to them and the temporal distribution over time in image plots and palette dot plots.

98

9 Spatial Variations, Identification and Analysis of Patterns and Textures

In general, colors are conceived as defined by the dimensions of hue, saturation and lightness. However, from the point of view of perception, there are many more factors that influence color appearance and the perception of colors correspondingly [Katz 1911] [Katz 1930] [Hurlbert 2013].

99

One of the most significant, but hitherto overlooked features of color appearance is *spatial variation*. By spatial variation we understand the change of hues related to spatial frequency in a given image. Such variations are related to several factors. Image complexity can be caused by cluttered image compositions with many small details, either in different or similar hues. Massive crowds of characters dressed in different colors are one type of subject that causes a high amount of spatial variation. Another type are layered image compositions with occlusion generated by objects in the foreground.

100

Visual complexity is connected to texture analysis in so far as spatial variations can be one feature that affects the legibility of image compositions (for a digital humanities approach to the investigation of image composition and style see Benini et al. 2016). An additional factor is the distribution of hues, with a high level of varying hues adding to visual complexity. At the same time an extremely uniform color distribution can lower legibility as well, if it is combined with a low degree of spatial variations and / or with darkness. Color separation and color attribution are a strong cue for object recognition and for scene detection [Hurlbert 2013] [Hansen and Gegenfurtner 2017].

101

In our aesthetic analyses the distinction between patterns and textures has been fundamental from the start, for several reasons. *Patterns* denote surface variations based on color attribution, for instance printed or woven patterns on fabrics, painted surfaces with patterns such as wallpapers etc. *Textures*, by contrast, refer to three-dimensional surface variations, such as knit-wear, rocks, brick walls, coarse unpolished wooden log structures. They invariably address tactile perception [Liu et al. 2015] [Zuo et al. 2016].

102

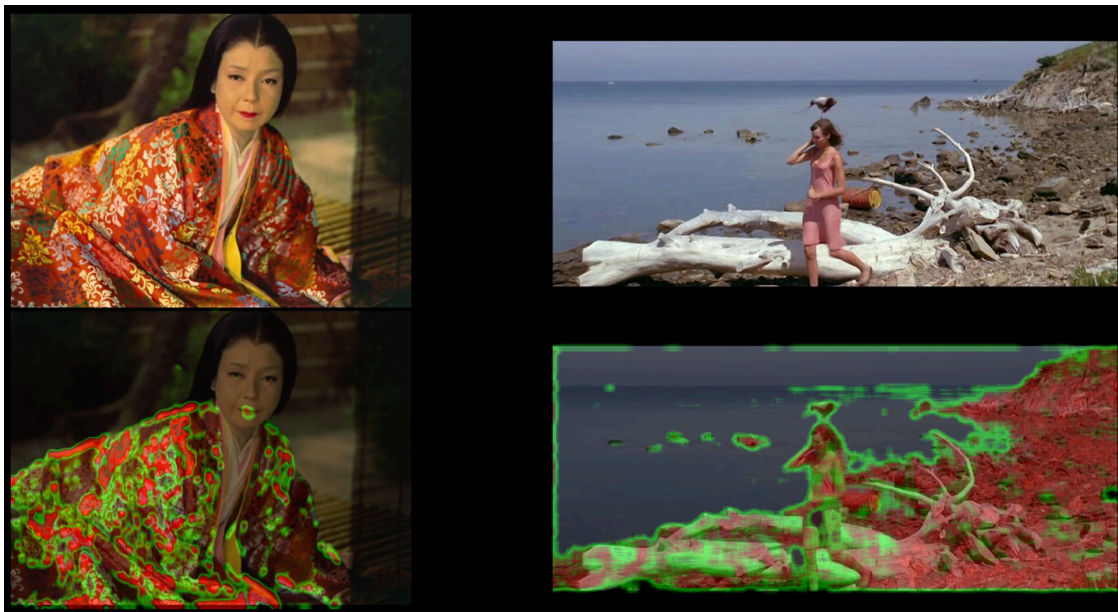


Figure 29. Patterns (left) in *Jigokumon* (JAP 1953, Teinosuke Kinugasa) vs. textures (right) in *Pierrot le fou* (FRA 1965, Jean-Luc Godard).

One guiding hypothesis of our research was a strong connection between the materiality of color film stocks and material properties of the *diegesis*, the spatio-temporal universe depicted in a film whose materials are selected and orchestrated by costume and production design. For instance half-tone printing as applied in Technicolor No. III to V dye-transfer processing is lacking definition due to problems to perfectly register the three printing dyes, which would be a prerequisite for spatial resolution of small-scale color variations. As a result we expected these films to omit patterns in their color design. Tinted films by contrast, lack spatial variation based on hues as they are uniformly colored by being submerged into dye baths, see *Timeline of Historical Film Colors* [Flueckiger 2012].

103

There is also a strong connection between affective modes of film perception and visual complexity or reduced legibility respectively. For instance in stressful scenes image complexity can increase substantially. Layered image compositions are one form to obstruct the automated perception of films that was regarded to be a cornerstone of the Hollywood system. As team member Michelle Beutler's research has shown, however, the Hollywood system itself was much less normative than previously assumed. The increase in tactile properties and affectively laden subjectivity noticed in the films of the 1960s onwards are at the center of Bregt Lameris's investigation on film colors and affect (see Lameris 2019). In Joëlle Kost's study of chromogenic film stocks visual complexity is one of the main topics as it relates to the improved resolution in these stocks.

104

By training a deep learning network to perform pixelwise sub-figure segmentation using the LIP dataset [Gong et al. 2017] we will

105

be able to analyze both features.

One possibility to assess the spatial frequency within a frame, is to use an edge detection algorithm, the intention is the more edges there are, the busier the region is. This has already shown to be a robust measure for spatial complexity, does however not cover solely hue and chroma related variance. VIAN currently visualizes three different measures as a heatmap over the player: The convolved edge density and the pixelwise luminance and $a*b$ channel variance.

106

9.1 Problems

Differentiation between patterns and texture computationally is a non-trivial task. A naïve approach would be to assume that variance in luminance tendentially indicates a tactile quality while high variance in hue and chroma would indicate patterns. However, since patterns are not excluded from high variance in the luminance channel, this approach does not yield accurate results. Furthermore, many materials that have a tactile quality for humans often do not differ significantly numerically from flat surfaces. Co-variance of spatial frequency, color values (hue, lightness, saturation) and textures vs. patterns is tightly connected to higher order processes in human visual perception, for instance color memory and cross-modal integration, i.e. the connection of tactile experience to visual and auditory perception. As shown previously in Flueckiger's investigation of sound design, material properties are often best detected by their acoustic cues. For a future, more elaborate system it would be an asset to include sound.

107

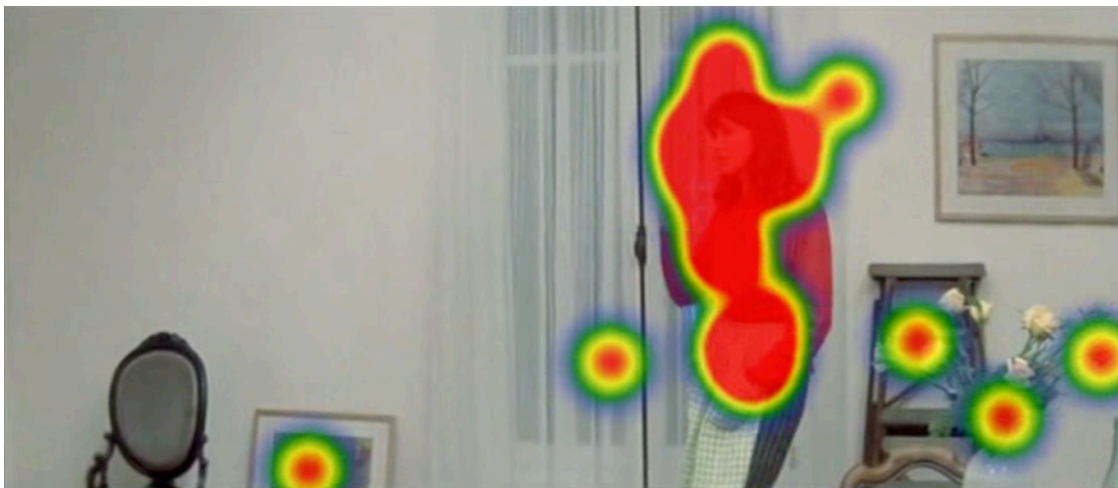
9.2 Lessons learned

Spatial frequency and the differentiation and assessment of patterns and textures are still part of our current research. Visual complexity is one of the most important factors when it comes to style and diachronic developments. Therefore we associated an eye-tracking study to the project to gain empirical insights into the topic (see Smith / Mital 2013; Rubo / Gamer 2018). The study was conducted by Miriam Loertscher in cooperation with Bregt Lameris. For this study we chose a set of exemplary scenes for different types of image composition and complexity, for instance the clear cut type without patterns and textures as in *Une femme est une femme*, the type "overwhelming object world" as in *Morte a Venezia* with completely cluttered, layered image compositions, or *Sayat Nova* (The Color of Pomegranates), a film that works with many textures and material variations, often by excluding the human figure.

108

Results are currently being processed, but from a brief look at the heatmaps, image parts with small-scale variations detract the viewers' gaze the most from the dominant focus on characters and most of all on faces. As Rubo and Gamer state "The influence of social stimuli and visual low-level saliency on eye movements have only recently been studied within the same datasets, and rarely in direct juxtaposition. During face perception, it was shown that facial regions diagnostic for emotional expressions received enhanced attention irrespective of their physical low-level saliency" [Rubo and Gamer 2018, 1]

109



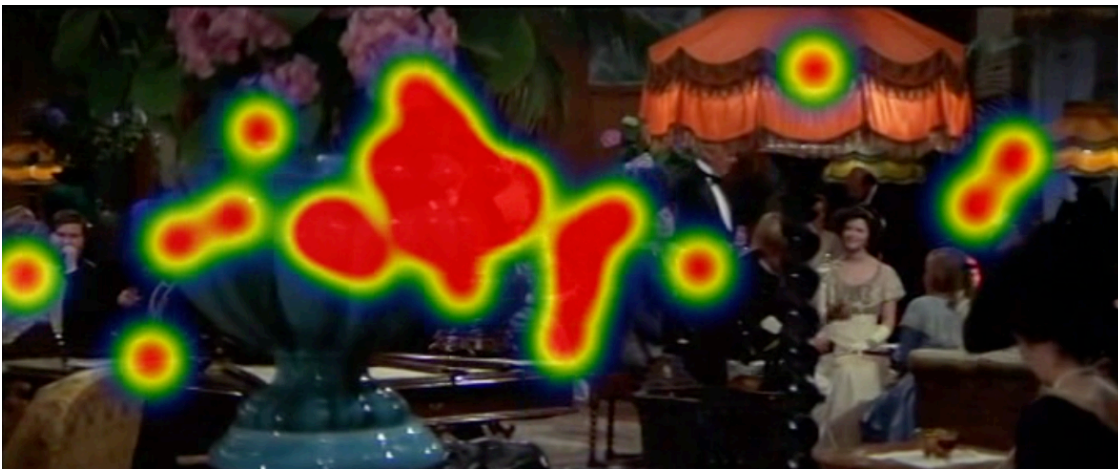


Figure 30. Heatmaps from the eye tracking study: Strong figure-ground separation and lack of spatial variation in the background in *Une femme est une femme* (above), textures 8 and patterns in combination with no humans in *Sayat Nova* (middle) plus the overwhelming object world with a high level of obstruction and visual complexity in *Morte a Venezia* (below). Study conducted by Miriam Loertscher and Bregt Lameris, ERC Advanced Grant *FilmColors*.

The resulting hypotheses have to be tested to identify regions of high spatial frequency and by comparison with the manually gathered data to assess if these regions represent pattern or textures.

110

10 Conclusion

In this article we discussed the potential and limitations of digital tools for the analysis of film aesthetics and narration based on the use case of research on the technology and aesthetics of film colors. Following the central argument established in the introduction, namely that such tools require a robust theoretical foundation, human interpretation, constant discussion and thorough reflection of the epistemological assumptions embedded in the tools, we explored various approaches to connect the humanities perspective with methods from data and computer science.

111

Our research has shown that we need to resort to a broad spectrum of finely grained analysis and visualization methods to avoid pitfalls of unfounded generalizations and anecdotal studies. On the downside of such a large dataset and sophisticated range of theoretical and analytical concepts there is a considerable amount of complexity and noise that tends to obscure clear-cut results. We found that for each of the research questions we need to take the full range of visualizations into account and re-evaluate the results on a case by case basis.

112

Compared to traditional, mostly language-dominated approaches to the aesthetics, technology and narratology of film colors, the digital humanities tools create evidence through the mapping of results into an instantly accessible array of visual representations. By relating detailed human annotation and interpretation to these visual representations, the integrated workflow consisting of the VIAN visual analysis software in combination with the crowdsourcing portal VIAN WebApp has created a comprehensive ecosystem for the investigation of film aesthetics and narration. Therefore it significantly extends established methods in film studies.

113

In currently running or planned cooperation projects we aim at exploring and extending this approach beyond the topic of film colors, in teaching, research and citizen science.

114

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement No 670446 *FilmColors*. Analyses were executed by the PhD candidates Olivia Kristina Stutz, Michelle Beutler, Joëlle Kost, PostDoc researcher Bregt Lameris, PI Barbara Flueckiger, and three student assistants Manuel Joller, Valentina Romero, Ursina Früh. Visualization and Multimedia Lab VMML at the University of Zurich directed by Renato Pajarola with Enrique Paredes and Rafael Ballester-Ripoll.

115

Notes

[1] <https://filmcolors.org/galleries/das-cabinet-des-dr-caligari-1919/>

Works Cited

- Arnold and Tilton 2020a** Arnold, Taylor; Tilton, Lauren (2020): "Enriching Historic Photography with Structured Data using Image Region Segmentation". In: *Proceedings of the First Artificial Intelligence for Historical Image Enrichment and Access (AI4HI)*. Marseille, France: Association for Computational Linguistics.
- Arnold and Tilton 2020b** Arnold, Taylor; Tilton, Lauren (2020 b): "Distant Viewing Toolkit. A Python Package for the Analysis of Visual Culture". In: *Journal of Open Source Software*, 5,45. *The Open Journal*, S. 1800, (<https://doi.org/10.21105/joss.01800>).
- Beau 2002** Beau, Frank (2002): "La solitude du technobole. Puissance politique des effets spéciaux". In: *CinémAction*, 102, pp. 196–207.
- Benini et al. 2016** Benini, Sergio; Svanera, Michele; Adami, Nicola; Leonardi, Riccardo; Kovács, András Bálint (2016): "Shot Scale Distribution in Art Films". In: *Multimedia Tools Appl.*, 75,23, Dez., pp. 16499–16527, (<https://doi.org/10.1007/s11042-016-3339-9>, accessed 11/17/2019).
- Bohn 2000** Bohn, Cornelia (2000): "Clothing as Medium of Communication". In: *Soziale Systeme*, 6,1, pp. 111–137.
- Bordwell 1989** Bordwell, David (1989): "Historical Poetics of Cinema". In: Barton Palmer (ed.): *The Cinematic Text. Methods and Approaches*. New York: ASM Press.
- Bordwell et al. 1985** Bordwell, David; Staiger, Janet; Thompson, Kristin (1985): *The Classical Hollywood Cinema. Film Style and Mode of Production to 1960*. London: Routledge.
- Brodbeck 2011** Brodbeck, Frederic (2011): "Cinematics. Film Data Visualization". In: *Cinematics*, (<http://cinematics.fredericbrodbeck.de/>, retrieved 05/30/2016).
- Burghardt et al. 2016** Burghardt, Manuel; Kao, M.; Wolff, C. (2016): "Beyond Shot Lengths. Using Language Data and Color Information as Additional Parameters for Quantitative Movie Analysis". In: Maciej Eder and Jan Rybicki (eds.): *Digital Humanities 2016. Conference Abstracts*. Kraków: Jagiellonian University & Pedagogical University, p. 753–755.
- Burghardt et al. 2017** Burghardt, Manuel; Hafner, Katharina; Edel, Laura; Kenaan, Sabrin-Leila; Wolff, Christian (2017): "An Information System for the Analysis of Color Distributions in MovieBarcodes". In: Maria Gäde (ed.): *Proceedings of the 15th International Symposium of Information Science (ISI 2017), Berlin, Germany, 13th-15th March 2017*. Glückstadt: Verlag Werner Hülsbusch, pp. 356–358, (<https://epub.uni-regensburg.de/35682/>, retrieved 09/08/2017).
- Casey and Williams 2014** Casey, Michael; Williams, Mark (2014): "Action. Audio-visual Cinematics Toolbox for Interaction, Organization, and Navigation of Film" (White Paper Report No. HD5139411), Hanover, Dartmouth College

(<https://hcommons.org/deposits/item/hc:12153/>, retrieved 07/12/2020)

Cutting 2016 Cutting, James E. (2016): *Perception, Attention, and the Structure of Hollywood Film*. (<http://people.psych.cornell.edu/~jec7/curreseach.htm>, retrieved 01/10/2018).

Cutting et al. 2012 Cutting, James E.; Brunick, Kaitlin L.; Candan, Ayse (2012): "Perceiving Event Dynamics and Parsing Hollywood Films". In: *Journal of Experimental Psychology*, Advance online publication, (<http://people.psych.cornell.edu/~jec7/pubs/jephppscenes.pdf>, retrieved 10/15/2016).

Dyer 2006 Dyer, Richard (2006): *Pastiche*. New York: Routledge.

ELAN ELAN. (<https://tla.mpi.nl/tools/tla-tools/elan/>, retrieved 10/15/2016).

Ferguson 2013 Ferguson, Kevin L. (2013): "Western Roundup". (<http://typecast.qwriting.qc.cuny.edu/2013/10/07/western-roundup/>, retrieved 07/11/2016).

Ferguson 2015 Ferguson, Kevin L. (2015): *Volumetric Cinema*. (<https://vimeo.com/119790662>, retrieved 11/07/2016).

Ferguson 2016 Ferguson, Kevin L. (2016): "The Slices of Cinema. Digital Surrealism as Research Strategy". In: Charles R. Acland and Eric Hoyt (eds.): *The Arclight Guidebook to Media History and the Digital Humanities*. Reframe Books, pp. 270–299, (<http://projectarclight.org/book/>).

Flueckiger 2011 Flueckiger, Barbara (2011): "Die Vermessung ästhetischer Erscheinungen". In: *Zeitschrift für Medienwissenschaft*, 5, pp. 44–60.

Flueckiger 2012 Flueckiger, Barbara (2012): *Timeline of Historical Film Colors*. (<http://filmcolors.org>).

Flueckiger 2017 Flueckiger, Barbara (2017): "A Digital Humanities Approach to Film Colors". In: *The Moving Image*, 17.2, pp. 71–94.

Flueckiger 2019 Flueckiger, Barbara (2019): "The Material Aesthetics of Len Lye's Experimental Color Films in the 1930s". Presentation *Len Lye Symposium*, Tinguely Museum Basel.

Flueckiger and Halter 2018 Flueckiger, Barbara; Halter, Gaudenz (2018): "Building a Crowdsourcing Platform for the Analysis of Film Colors". In: *The Moving Image*, 18.1, pp. 80–83.

Flueckiger et al. 2017 Flueckiger, Barbara; Evirgen, Noyan; Paredes, Enrique G.; Ballester-Ripoll, Rafael; Pajarola, Renato (2017): "Deep Learning Tools for Foreground-Aware Analysis of Film Colors". In: *AVinDH SIG*, Digital Humanities Conference Montreal.

Genette 1972 Genette, Gérard (1972): *Figures III*. Paris: Seuil.

Genette 1983 Genette, Gerard (1983): *Nouveau Discours du récit*. Paris: Le Seuil.

Genette 1992 Genette, Gérard (1992): *Palimpsestes. La littérature au second degré*. Paris, Seuil.

Giunti 2010 Giunti, Livia (2010): *Problemi dell'analisi del testo di finzione audiovisivo. Verifica e sviluppo di un modello analitico e interpretativo con strumenti digitali*. Università degli Studi di Pisa.

Giunti 2014 Giunti, Livia (2014): "L'analyse du film à l'ère numérique. Annotation, geste analytique et lecture active". In: *Cinéma & Cie*, 14,22/23, pp. 127–143.

Gong et al. 2017 Gong, Ke; Liang, Xiaodan; Zhang, Dongyu.; Shen, Xiaohui; Lin, Liang (2017): "Look into Person. Self-supervised Structure-sensitive Learning and A New Benchmark for Human Parsing." In: arXiv:1703.05446, (2017). (<http://arxiv.org/abs/1703.05446>, retrieved 07/12/2020)

Gruber 2009 Gruber, Klemens; Wurm, Barbara; Kropf, Vera (eds.) (2009): *Digital Formalism. Die kalkulierten Bilder des Dziga Vertov*. Wien: Böhlau Verlag.

Hagberg et al. 2008 Hagberg, Aric A.; Schult, Daniel A.; Swart, Pieter J. (2008): "Exploring Network Structure, Dynamics, and Function using NetworkX". In: Gaël Varoquaux, Travis Vaught and Jarrod Millman (eds.): *Proceedings of the 7th Python in Science Conference*. Pasadena, CA USA, pp. 11–15.

Hahn 2009 Hahn, Stefan (2009): "Filmprotokoll Revisited. Ground Truth in Digital Formalism". In: Klemens Gruber, Barbara Wurm and Vera Kropf (eds.): *Digital Formalism: Die kalkulierten Bilder des Dziga Vertov*. Wien: Böhlau Verlag, pp. 129–136.

Halter et al. 2019 Halter, Gaudenz; Ballester-Ripoll, Rafael; Flueckiger, Barbara; Pajarola, Renato (2019): "VIAN. A Visual Annotation Tool for Film Analysis". In: *Computer Graphics Forum*, 38,3, pp. 119–129.

Hansen and Gegenfurtner 2017 Hansen, Thorsten; Gegenfurtner, Karl R. (2017): "Color contributes to object-contour perception in natural scenes". In: *Journal of Vision*, 17,3, März, pp. 14–14.

Heftberger 2015 Heftberger, Adelheid (2015): "'Die Verschmelzung von Wissenschaft und Filmchronik'. Das Potenzial der Reduktionslosen Visualisierung am Beispiel von Das Elfte Jahr und Der Mann mit der Kamera von Dziga Vertov". In: *La Visualisation des Données en Histoire = Visualisierung von Daten in der Geschichtswissenschaft*, pp. 229–263.

- Heftberger 2016** Heftberger, Adelheid (2016): *Kollision der Kader. Dziga Vertovs Filme, die Visualisierung ihrer Strukturen und die Digital Humanities*. München: Edition Text + Kritik. [English translation: Heftberger, Adelheid (2018): *Digital Humanities and Film Studies. Visualising Dziga Vertov's Work*. Berlin: Springer International Publishing.]
- Hurlbert 2013** Hurlbert, Anya (2013): "The Perceptual Quality of Color". In: Liliana Albertazzi (ed.): *Handbook of Experimental Phenomenology. Visual Perception of Shape, Space and Appearance*. New York: John Wiley & Sons, Ltd SN, pp. 369–394.
- Itten 1970** Itten, Johannes (1970): *Kunst der Farbe*. Ravensburg: Ravensburger Buchverlag.
- Jameson 1991** Jameson, Fredric (1991): *Postmodernism. Or, the Cultural Logic of Late Capitalism*. London: Verso.
- Katz 1911** Katz, David (1911): *Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung*. London: Kegan Paul, Trench, Trubner & Co. Ltd.
- Katz 1930** Katz, David (1930): *Der Aufbau der Farbwelt*. Leipzig: Johann Ambrosius Barth.
- Lameris 2019** Lameris, Bregt (2019): "Hallucinating Colours. Psychedelic Film, Technology, Aesthetics and Affect". In: *Cinéma & Cie. Special Issue Cinema and Mid-Century Colour Culture*, 32, Spring.
- Ledig and Ullmann 1988** Ledig, Elfriede; Ullmann, G. erhard (1988): "Rot wie Feuer, Leidenschaft, Genie, Wahnsinn. Zu einigen Aspekten der Farbe im Stummfilm". In Elfriede Ledig (ed.): *Der Stummfilm. Konstruktion und Rekonstruktion*. München, Schaudig, Bauer, Ledig, pp. 89–116.
- Liu et al. 2015** Liu, Jianli; Lughofer, Edwin; Zeng, Xianyi (2015): "Aesthetic Perception of Visual Textures. A Holistic Exploration Using Texture Analysis, Psychological Experiment, and Perception Modeling". In: *Frontiers in Computational Neuroscience*, 9, p. 134.
- Long et al. 2015** Long, Jonathan; Shelhamer, Evan; Darrell, Trevor (2015): "Fully Convolutional Networks for Semantic Segmentation". pp. 3431–3440, (<https://arxiv.org/abs/1411.4038>, retrieved 07/12/2020).
- Manovich 2012** Manovich, Lev (2012): "How to Compare One Million Images?" In: D. Berry (ed.): *Understanding Digital Humanities*. London: Palgrave Macmillan UK, pp. 249–278.
- Manovich 2013a** Manovich, Lev (2013): "Kino-Eye in Reverse. Visualizing Cinema". In: Jeffrey Geiger and Karin Littau (eds.): *Cinematicity in Media History*. Edinburgh: Edinburgh University Press, pp. 211–234.
- Manovich 2013b** Manovich, Lev (2013): "Visualizing Vertov". In: *Russian Journal of Communication*, 5,1, pp. 44–55.
- Mazzanti 1998** Mazzanti, Nicola (1998): "The Colours of the Film d'Arte Italiana". In: Luciano Berriatúa, et al., *Tutti i colori del mondo. Il colore nei mass media tra 1900 e 1930. = All the colours of the world*. Reggio Emilia, Edizioni Diabasis, pp. 141–146.
- Mazzanti 2009** Mazzanti, Nicola (1998): "Colours, Audiences, and (Dis)Continuity in the 'Cinema of the Second Period'". In: *Film History*, 21,1, pp. 67–93.
- Melgar et al. 2017** Melgar Estrada, Liliana; Hielscher, Eva; Koolen, Marijn; Olesen, Christian Gosvig; Noordegraaf, Julia; Blom, Jaap (2017): "Film Analysis as Annotation. Exploring Current Tools". In: *The Moving Image: The Journal of the Association of Moving Image Archivists*, 17,2, pp. 40–70.
- Olesen 2017** Olesen, Christian Gosvig (2017): *Film History in the Making. Film Historiography, Digitised Archives and Digital Research Dispositifs*. Amsterdam: University of Amsterdam, (<https://dare.uva.nl/search?identifier=ad68a275-e968-4fce-b91e-4783cd69686c>, retrieved 10/07/2017).
- Olesen et al. 2016a** Olesen, Christian Gosvig; Gorp, Jasmijn van; Fossati, Giovanna (2016): "Datasets and Colour Visualizations for 'Data-Driven Film History. A Demonstrator of EYE's Jean Desmet Collection'". In: *Creative Amsterdam. An E-Humanities Perspective. A Research Program at the University of Amsterdam*, (<http://www.create.humanities.uva.nl/results/desmetdatasets/>, retrieved 11/11/2016).
- Olesen et al. 2016b** Olesen, Christian Gosvig; Masson, Eef; Gorp, Jasmijn van; Fossati, Giovanna; Noordegraaf, Julia (2016): Data-Driven Research for Film History. Exploring the Jean Desmet Collection. In: *The Moving Image*, 16,1, (<https://muse.jhu.edu/article/640569>).
- Pasue and Walkowski 2018** Pause, Johannes; Walkowski, Niels-Oliver (2018): "The Colorized Dead. Computerunterstützte Analysen der Farblichkeit von Filmen in den Digital Humanities am Beispiel von Zombiefilmen", (<https://edoc.bbaw.de/frontdoor/index/index/docId/2591>).
- Plantinga 2009** Plantinga, Carl (2009): *Moving Viewers. American Film and the Spectator's Experience*. Berkeley: University of California Press.
- Redmon et al. 2013** Redmon, Joseph; Divvala, Santosh; Girshick, Ross; Farhadi, Ali (2015): You Only Look Once. Unified, Real-Time Object Detection. In: *arXiv:1506.02640 [cs]*, Jun., (<http://arxiv.org/abs/1506.02640>, retrieved 05/29/2017).
- Reyes-García 2013** Reyes-García, Everardo (2013): "On Visual Features and Artistic Digital Images". New York: ACM, (<http://dl.acm.org/citation.cfm?id=2466835>).
- Reyes-García 2014** Reyes-García, Everardo (2014): *Explorations in Media Visualization*. New York: ACM,

(http://www.academia.edu/download/35860006/Reyes_2014-Explorations_in_Media_Visualization.pdf,%20%20%20%20%20%20%20%20%20%20%20%20%20http://ceur-ws.org/Vol-1210/datawiz2014_11.pdf, retrieved 07/12/2020).

Reyes-García 2017 Reyes-García, Everardo (2017): *The Image-interface. Graphical Supports for Visual Information*. Hoboken, NJ: Wiley-ISTE.

Reyes-García and Bouhai 2017 Reyes-García, Everardo; Bouhai, Nasreddine (2017): *Designing Interactive Hypermedia Systems*. In: Nasreddine Bouhai (ed.). Wiley-ISTE.

Rosebrock 2014 Rosebrock, Adrian (2014): "How-To. OpenCV and Python K-Means Color Clustering". In: *PyImageSearch*, (<http://www.pyimagesearch.com/2014/05/26/opencv-python-k-means-color-clustering/>, retrieved 06/26/2017).

Ross 2007 Ross, Jacqui (2007): *Colour Analysis Tools in ImageJ*. (https://www.fmhs.auckland.ac.nz/assets/fmhs/sms/biru/docs/Colour_Analysis_Tools_in_ImageJ.pdf, retrieved 0/12/2020).

Rother et al. 2004 Rother, Carsten; Kolmogorov, Vladimir; Blake, Andrew (2004): "GrabCut. Interactive Foreground Extraction using Iterated Graph Cuts". In: *ACM Transactions on Graphics (SIGGRAPH)*, Aug.

Rubo and Gamer 2018 Rubo, M. and Gamer, M. "Social Content and Emotional Valence Modulate Gaze Fixations in Dynamic Scenes", *Scientific Reports*, 8.1 (2018): 1–11. (<https://doi.org/10.1038/s41598-018-22127-w>, retrieved 07/12/2020)

Smith 1995 Smith, Murray (1995): *Engaging Characters. Fiction, Emotion, and the Cinema*. Oxford, New York: Oxford University Press.

Smith and Mital 2013 Smith, Tim J.; Mital, Parag K. (2013): "Attentional Synchrony and the Influence of Viewing Task on Gaze Behavior in Static and Dynamic Scenes". In: *Journal of Vision*, 13,8, Juli, S. 16–16, (<https://jov.arvojournals.org/article.aspx?articleid=2193975>, abgerufen 03/31/2020).

Stutz 2016 Stutz, Olivia Kristina (2016): *Algorithmische Farbfilmästhetik. Historische sowie experimentell-digitale Notations- und Visualisierungssysteme des Farbfilms im Zeichen der Digital Humanities 2.0 und 3.0*. Zürich: Universität Zürich, (http://www.film.uzh.ch/dam/jcr:bed543b6-4a67-4ff8-8f51-b85a739417d5/MA_AlgorithmischeFarbfilmaesthetik_Stutz_HS2016_def.pdf, retrieved 04/09/2017).

Tan 1996 Tan, Ed (1996): *Emotion and the Structure of Narrative Film. Film as an Emotion Machine*. Mahwah, N.J, Lawrence Erlbaum Assoc Inc (1996).

Thompson 1986 Thompson, Kristin (1986): "The Concept of Cinematic Excess". In: Philip Rosen, Leo Braudy and Marshall Cohen (eds.): *Film Theory and Criticism. Introductory Readings*. New York: Columbia University Press, pp. 487–498.

Thompson 1988 Thompson, Kristin (1988): *Breaking the Glass Armor*. New Jersey: Princeton University Press.

Vänskä 2017 Vänskä, Annamari (2017): "Gender and Sexuality". In: Alexandra Palmer (ed.): *A Cultural History of Dress and Fashion*. Bloomsbury Academic, pp. 107–129.

Wilkening 2014 Wilkening, Anke (2014): "Die Restaurierung von Das Cabinet des Dr. Caligari". In: *VDR Beiträge zur Erhaltung von Kunst- und Kulturgut*, 2, pp. 27–47.

Zhao et al. 2016 Zhao, Hengshuang; Shi, Jianping; Qi, Xiaojuan; Wang, Xiaogang; Jia, Jiaya (2016): "Pyramid Scene Parsing Network". In: *arXiv:1612.01105 [cs]*, Dec., (<http://arxiv.org/abs/1612.01105>, retrieved 03/27/2018).

Zuo et al. 2016 Zuo, Hengfeng; Jones, Mark; Hope, Tony; Jones, Robin (2016): "Sensory Perception of Material Texture in Consumer Products". In: *The Design Journal*, 19.03, pp. 405–427.