Materiality and Montage: Film Studies, Digital Humanities and the Visualization of Moving Images

Digital Humanities and Film Studies

In this paper, I will highlight some recent initiatives in the study of film within the digital humanities, in which context I will also present some of my own endeavors, specifically visualizations created in collaboration with the pioneering new media theorist Lev Manovich from films made by the Soviet avant-garde director Dziga Vertov (1896-1954). Following this, I will discuss some of the issues related to the use of visualizations as an aid to scholarly research. Finally, I will address a number of possible research questions in film and media studies, answers to which may benefit significantly from the collaboration between film/media scholars and computer scientists on the one hand, and (moving image) archivists on the other.

Before proceeding to discuss the situation in which film archives currently find themselves, I would like to share a few introductory thoughts from my own perspective as a practitioner in digital humanities. As it has become good scholarly practice to begin by stating one's own working definition of the term, I would like to offer the following quote by Eric Hoyt, Kit Hughes and Charles R. Acland (2016, 3), which appeals to me for its integrative approach to digital humanities:

> Rather than take digital humanities as a circumscribed field of research, pedagogy, and outreach, we understand DH as a strategically deployed term of mutual recognition that enables contemporary knowledge workers to signal a shared project interested in the relationship between digital technologies and humanities work. Of course, disagreement exists over what that project is. In a sense, we are all digital humanists. Article databases, online catalogues, search algorithms, word processing software, email, and course management systems already shape contemporary academic work in countless ways.

Their definition would appear to become increasingly valid when it comes to the arguably somewhat conservative views of film archives and film museums or even cultural heritage institutions altogether. Recent discussions within the archival community have shown a tendency to be quite emotionally charged and strongly motivated by personal and national interests when it comes to the preservation of analog film, while for the most part the outside world does not seem to understand what the issue at stake really is. Film and media studies, as academic disciplines, are by their very nature highly affected by the digital (media) turn and its effect in modifying their primary material of study. According to the American Heritage Dictionary, we can define a medium as a specific kind of artistic technique or means of expression as determined by the materials used or the creative methods involved. Whether or not the material properties of the original carrier is relevant to a particular study, depends very much on the research focus. In some cases, direct engagement with the original carrier, for example a 35mm film print, is still preferred, while increasingly video formats or digital files tend to suffice. While film historians and archivists have a tendency to view the "original" as the only legitimate source, most film scholars seem quite content with being able to view the content regardless of its format and quality

In digital humanities, the question of high quality source material has once again become relevant. There is nothing wrong per se with using files extracted from a commercial DVD, about which one tends to have very little knowledge when it comes to the materials and processes involved in their creation. However, for most of the examples I will refer to it has proved absolutely necessary to work with digital copies made from reliable sources under carefully controlled conditions. This is where the successful collaboration between film scholars and film archives becomes vital for innovative and serious research. Admittedly, the level of success still depends largely on personal relations or lucky circumstances. Nonetheless, I would like to argue for a more formalized working relationship between GLAMs and research institutions so that reliable archival documents can be made available for further research in sufficient quality. There are, of course, a number of obstacles, not least the issue of quality, which cannot presently be guaranteed by archives.

However, friction does not only occur between archivists and film/media scholars, but amongst the scholars themselves, for example between different disciplines such as humanities and computer sciences. For the author and scholar of digital literature Mark C. Marino[1] the traditionally ascribed roles simply do not work anymore. He argues (Kudenov 2016) that the difference between the disciplines does not lie merely in the degree to which they engage in interpretation and creativity:

> I've changed my way of entering the realm of code. I've learned how to be a guest in someone else's domain, how to respect the interpretive work that computer scientists already do, and how to see through the artificial separation between the so-called two worlds. A lot of what computer scientists already do involves if not interpretation then interacting with software as a mode of discourse. In other words, in order to speak code, one needs to understand its registers of meaning in to achieve that one must spend time chatting with those who speak code as primary language.

In the same way, computer scientists like Matthias Zeppelzauer (Olesen n.Y.) may gain new insights from working with film archivists on interdisciplinary projects: "I remember that we had numerous discussions with our colleagues from the Austrian Film Museum about the state of the material, the artifacts and the challenges that originate from them for automatic analysis. Based on these discussions and the inspection of the material we developed a comprehensive understanding of this specific type of film material." Zeppelzauer later points out that knowledge exchange in this case worked both ways: "As an additional output of our discussions the film archivists developed a better understanding of how a computer 'sees' images and videos and why particular artifacts which are to a wide degree compensated by a human viewer are highly disturbing in automatic film analysis (e.g. shaking and flicker)." Lev Manovich (2012, 473) describes the typical model in place as a dependence of humanists on computer scientists for processes such as data mining. His vision for the future is to follow a different strategy, however: "We want humanists to be able to use data analysis and visualization software in their daily work, so they can combine quantitative and qualitative approaches in all their work. How to make this happen is one of the key questions for the digital humanities". Nick Redfern (2013) has argued along similar lines, even though he does not talk about digital humanities per se, but rather the need for a so-called statistical literacy in film studies: "Along with many other disciplines in the humanities, Film Studies has simply failed to grasp the importance of statistical

1 Marino has also written interactive children's stories. For more information on the project "Mrs. Wobble and The Tangerine" see: http://markcmarino.com/mrsw/

literacy to everyday life, to students' employability, and to the specific demands of the discipline." This also means providing film scholars with the necessary tools for carrying out certain comparably simple procedures normally assigned to computer scientists: shot recognition, image recognition, speech-to-text recognition, as well as the possibility to annotate films collectively. However, video annotation is still in its early stages as far as a methodology and software for humanities' users are concerned. The most common tools for film studies still seem to be free software like ANVIL[2] or ELAN[3]. According to Niels-Oliver Walkowski (2016), the topic of annotating artworks continues to raise significant interest in the digital humanities and has become an increasingly complex issue due to the development of computational environments, the usage scenarios and interpretation. Within the DARIAH network a workgroup for "Digital Video Annotation" has now been set up and will work toward best practices and knowledge exchange.

One can state with a certain degree of confidence that not a lot of research has been done within the digital humanities when it comes to audiovisual media, especially when we compare to the efforts made in text analysis.[4] There are a number of different reasons for this. A "film" is essentially a multimedia, collaborative and performative artwork that encompasses image, audio, text as well as a range of other documents, mostly ephemeral in nature, which are linked to its production, promotion and/or presentation. Scholars therefore struggle with an abundance of sources, particularly if they are seriously interested in engaging with film in a more comprehensive and data-driven way. At the same time access to primary sources is in many cases still lacking, and the databases currently employed by film archives and other collecting institutions as well as online sources tend to differ, at times substantially, in terms of metadata quality and the available search functions.

While digital humanities is still very much text-oriented, interest in moving image media is slowly but steadily emerging. Of the somewhat disparate groups and individual researchers working on moving images within a digital humanities context, I would like to highlight only a few initiatives here. Two of the pioneers when it comes to combining film studies and digital humanities are Lev Manovich, founder and director of the Cultural Analytics Lab, and Jeffrey Schnapp, faculty director of metaLAB at Harvard. A well-established collaborative online project for quantitative film analysis is "Cinemetrics," founded by Yuri Tsivian (www.cinemetrics.lv), which consists to a large degree of manually produced data on the average shot lengths of films. The website unquestionably functions as a hub for all those interested in exploring formal film analysis and participating in the network. This pioneering project was set up ten years ago with a high degree of personal dedication, and it is constantly being updated with new features.[5] The data is available freely for others who wish to run

---

2 For more information see: http://www.anvil-software.org/

3 For more information see: https://tla.mpi.nl/tools/tla-tools/elan/

4 Some scholars try to provide information on film studies and digital humanities, such as Catherine Grant in her pioneering project "Film studies for free" (http://filmstudiesforfree.blogspot.de/), or the journal "[in]Transition" (http://mediacommons.futureofthebook.org/intransition/), where videographic film studies are explored. There is also the "Special Interest Group in Audiovisual Materials in Digital Humanities" (https://avindhsig.wordpress.com/) and a collection of film-related projects in the digital humanities available on the website of the Transformations Conference Blog (https://transformationsconference.net/dh-cinema-projects/).

5 For example, Cinemetrics labs was created lately with exactly that reason in mind and is a function worth exploring. For more information see: http://www.cinemetrics.lv/labs.php. Barry Salt, one of the pioneers of quantitative film analysis whose work over the last decades has contributed in a significant way to Cinemetrics, should also not go unmentioned.

their own statistical analyzes. However, when it comes to Cinemetrics, one needs to keep the human factor in mind, as the measurements are not always entirely accurate and the metadata provided by the contributors could sometimes be greatly improved. At other times, it is not easy to determine which version, or even which parts, of a particular film were measured, and thereby impeding the comparison of filmic structures with other films, which in my view would be one of the biggest advantages of the tool. The project "Mapping Desmet", meanwhile, investigates and visualizes the screening history of the famous Desmet collection preserved at the EYE Film Institute Netherlands.[6] Another ambitious historical film initiative has been set by the film scholar Radomir Kokeš, whose focus is on early Czech film (http://www.douglaskokes.cz/pdz/), while the computer scientist Manuel Burghardt (2016) concentrates on the analysis and visualization of film color[7] and film dialogue.

Additionally, a handful of useful websites have emerged recently which collect information about digital tools for scholarly research, for example the Directory of Digital Research Tools (DiRT).[8] A comprehensive list of tools, software and related projects, particularly as pertains to data visualization, can also be consulted on the Cultural Analytics Lab's website.[9] Another noteworthy example is the recently published collection of papers *The Arclight Guidebook to Media History and the Digital Humanities*. The Arclight project, which was supported by a Digging into Data grant from the U.S.'s Institute for Museum and Library Services and Canada's Social Sciences and Humanities Research Council, is carefully explained by Eric Hoyt in his video essay "Data Mining Silent Cinema History" (https://www.youtube.com/watch?v=qO8W_ccIX7Y). However, while Hoyt (2014) talks about applying methods from text analysis like topic modeling to film historical research, he unfortunately does not elaborate on the potential of audiovisual analysis but rather remains firmly entrenched in text-based data mining and the analysis and visualization of film magazines.[10]

As hopefully becomes clear, an increasing number of resources are being made available and many individuals are beginning to work on innovative projects which fall under the umbrella of film studies in digital humanities, but the network is only forming slowly. If we want to shape digital humanities into a genuine dialogue, we need to work together on methodology, not just attend summer schools, as Scott Weingart (2016) writes: "We need to make room in our curricula for actual methods courses, or even degrees focused on methodology, in the same fashion as social scientists, if we want to start a robust practice of developing appropriate tools for our own research." Jeffrey Schnapp has used the phrase "Knowledge Design" as an overarching concept which in my view encompasses disciplines as well as institutions. He views the current situation in the humanities as one of experimentation rather than of using clear-cut methods. According to Schnapp (2016, 6), there are new challenges arising, such as "how to construct arguments that zoom back and forth between the micro, the meso, and the macro, perhaps even overleaping those middle layers of analysis and

---

6 For more information see: http://mappingdesmet.humanities.uva.nl/#/. Project leader Christian Gosvig Olesen has also collected a bibliography available here: https://filmhistoryinthemaking.com/digital-film-historiography-a-bibliography/.

7 Other projects working on color charts and comparisons include "Moviebarcode" (http://www.redbubble.com/de/people/moviebarcode/portfolio).

8 For more information see: http://dirtdirectory.org/

9 Formerly Software Studies Initiative. For more information see: http://www.culturalanalytics.info/

10 For more information see: http://search.projectarclight.org/. Together with David Pierce, Hoyt is also responsible for the Lantern website: http://lantern.mediahist.org/

narrative that once constituted the home turf of the arts and humanities disciplines?" As the main nodes for entry points he clusters into the following fields of activities or concepts (ibid.): storied collections (innovative ways of working with and across collections), social lives of things (multimedia approaches to the description and representation of three-dimensional objects as networks of relations), new learning containers (rethinking learning spaces and models), and ubiquitous curation (the world as laboratory).

Consequently this would call for collaborative efforts across institutional and disciplinary borders rather than trying to maintain some kind of hegemony on methodology. This paper too is driven by the belief that film studies in digital humanities has to be a collaborative and interdisciplinary undertaking which crosses institutional borders, placing archivists and curators on an equal footing with academics and researchers by having the latter group acknowledge the expertise and scholarly activities of the former. This can hardly be considered a revolutionary concept and should not come as anything of a surprise, but is still a fairly rare occurrence in practice and therefore bears explicit repeating. By presenting some examples of my own work as both an archivist as well as a film and media scholar I can hopefully reveal just some of the fascinating results that can be achieved from this kind of collaboration.

Investigations into Film Style

Any film is an artwork which has an inherently modular character, yet its parts (e.g. audio, video, text) are deeply coherent. In addition, film has always been essentially a collective endeavor, geared towards mass entertainment and distribution on international markets. Film and TV analysis provides a firmly established set of methods for the formal description and analysis of audiovisual material, Generally, it aims at the analysis of distinct formal elements and their functional occurrence to determine key artistic devices or certain patterns as well as the meaning of interconnected visual motifs. This can be dated back to the Russian formalist school, which drew upon the formal characteristics of an artwork and can arguably be viewed as the most systematic and interdisciplinary approach to date. Also the Prague linguistic circle with its particular focus on poetology contributed to developing the methodology. In the following chapter, I will outline the potential of quantitative analysis (manual or automatic) and visualization for film and media studies, particularly for film history. This can work on the premise that a film represents a kind of database of formal elements, from which distinct temporal or thematic entities can be extracted, e.g. shot lengths, scene changes, shot composition and scale, visual motifs and camera movement. Historical attempts to "measure" films in this way date back to the beginning of the medium and appear regularly again throughout its history.

Messages expressed through films, whether a propaganda campaign for the Soviet State or an individual aesthetic conviction, are conveyed in one way or another by formal devices. Because of this, the arguments which are used to transfer these messages can be formalized to some degree and represented visually. That said, there is obviously no standard schema or template that applies to all the different forms of communication. Johanna Drucker (2014, 22) argues along much the same lines, stating that, "[t]he workings of power, the force of ideology, the transmission of values, and other

abstract ideas have no specific visual form, even if they work through a material social world." However, as has been investigated within the visual studies and communication studies disciplines, systematic uses of visual images have created de facto standards. Knowing the context is essential: Is our reference point the Soviet Union of the 1930s or the Austrian rural community of the 1950s? In other words, research into visual codes is both possible and indeed necessary, especially as concerns artworks which were constructed with a specific plan or purpose in mind. Artistic choices convey meaning, as Drucker (ibid., 31) writes, by, "the force of diagonals, emotive qualities of color, or other formal features." The meanings behind certain styles, motifs, textures, colors etc. can only be mined, visualized and understood if humanities scholars have already provided the basis upon which to choose and subsequently analyze visual and formal entities. Why not try to create digital tools which automatically analyze large data sets according to their formal characteristics and support a visual history based on the chosen parameters and settings?[11]

One possible field of investigation which immediately comes to mind is the analysis of film style. Early film theorists like Béla Balázs (1982 [1925]) and the key figures of the aforementioned Russian formalist school in particular, Viktor Šklovskij (1984 [1925]), Boris Kazanskij (2005 [1927]), Semen Timošenko (1928) and Boris Ėjchenbaum (2005 [1927]), initiated studies centered on questions of style in film and literature. According to Šklovskij (1984 [1925], 144), the content of a literary work in formalist understanding consists of the sum of all the stylistic devices used, while choices concerning the likes of material and form are due to a certain construction principle. Thus, the precise organization of the shooting schedule and actors is indispensable, although, as Kazanskij reminds us (2005 [1927], 125), sometimes the best moments in a film still occur by pure chance. Kazanskij highlights devices like composition, lighting, camera angle and shot scale, which can later on be woven into a film's style. In the end, however, it is the montage, he writes (ibid., 123), which serves as a film's intrinsic stylistic device: while everyone from cameraman to the actor to the set designer contributes in some way to the creation of the images, the mastery of film lies only in the montage of those images. The work of the Formalists was revived and expanded in the 1980s by film scholars like David Bordwell, Janet Staiger and Kristin Thompson (1985) or Barry Salt (1992). Of particular interest to them were topics such as the notion of national film styles, Hollywood studio styles or early film history in general.

The ability to compare large corpora of film data now can influence and shape our understanding of film style further and allows us to make comparisons between protagonists in the film industry, like directors or actors, far more easily. It might also facilitate synchronous as well as diachronic analysis. How can, for example, the changes in film editing style during the transition from silent to sound film in the late 1920s and early 1930s be described, formalized and visualized, as scholars like Charles O'Brien (2005) or Lea Jacobs (2014) have already attempted to carve out using smaller samples?

Another area of application would be to utilize software for the identification of undated or unidentified films which abound in archival collections. Could film style patterns which have previously been sampled or calculated, e.g. for certain time periods, serve as a useful parameter for retrieval

11 As a historical example, Drucker refers us to Walter Cranes *Line and Form* (1900), one of the attempts to formulate universal visual patterns.

tools? This idea has been brought up already in the 1980s by the German film scholar Helmut Birett, who, although a geologist by training, has become a reference for German film history. He suggested to apply additional statistical methods in order to help identify films; a nigh on impossible or at best extremely time consuming task, given that around 80% of all silent films are alleged to be lost. Birett's approach is a comparative one that attempts to locate parameters which can be cross referenced, like the overall length of a film as well as individual shot lengths and intertitles. Thus, if the corpus is large enough, we can define what he calls the "Normalfilm" (average film). Any deviations can be analyzed separately, and unidentified films can subsequently be checked for correlations. Additionally, we might be able to follow the development of a director's technique over time, and once we have defined his "Normalfilm", we can, according to Birett (1988), go on to discuss his stylistic characteristics. Although yet to materialize, such ideas would in my view be worthy of further investigation on a larger scale.

Birett raises the important question of how films can be compared and answers it by saying that we should focus on temporal units rather than narrative. While Cinemetrics as well as its forefather Barry Salt don't explicitly refer to Birett, they can be seen to follow his view to a certain degree. Measuring films might be able to support hypotheses about certain genres or the history of style in different cinematic cultures (like Germany, Scandinavia, USA or Russia). Was there a gradual evolution of style or rather radical changes? Are manifestations of sociopolitical influences detectable, like censorship decisions or changes in power? When and where did the major changes occur and how are we able to explain these deviations from the usual pattern? How can we combine measuring shot lengths (or image composition etc.) with methods of qualitative analysis? Which topics can be investigated in films by taking formal analysis into account? In this context, however, it cannot be stressed enough that a great many written, photographic and audiovisual documents have yet to be digitized, and one therefore has to be very careful with assumptions about stylistic features.

Film historians are often interested in answering questions about the local, national or international circulation and reception of films. Where and how often was a particular film screened, and which other films were scheduled in the same screening program? The research questions can be less focused on the history of the film per se, but rather on its audience. How did the audience react? Do we know anything about the social composition of the viewers? How was the film received? Who wrote about it? One line of inquiry would be to consult all the reviews concerning one film at a specific time and place. How can I search these documents for recurring topics of interest related to the film, for example war, housing problems, poverty, political issues? Here, of course, it would be useful to be able to include not only journal articles but also images, video content, private or semi-private blogs, comments, etc.

Finally, if I as a researcher am more interested in visual motifs or image composition, it would be helpful for me to find out (or even visualize) how one film(maker) influenced others or how the same motifs alter over time, and how and why certain scenes and images are re-used in later films (e.g. documentaries). Re-use and intertextuality are just two keywords to mention briefly while pointing out that images develop lives of their own and change their meaning in different contexts. This is basic montage theory as laid down by the pioneers such as Vsevolod Pudovkin and Lev Kulešov in the 1920s and 1930s. We could also trace the origins of iconic images and ascertain how

they became part of our collective memory, such as the countdown to the launch of the rocket ship in Fritz Lang's film "Frau im Mond" (1929). Research into film style can therefore be viewed as a process by which to monitor the trajectories of visual images throughout film history and analyze their different meanings.

The Visualization of Time-Based Media

In the previous chapter, I had outlined some of the ways one can formalize and quantify film works. I will now focus on depicting the information one has gathered in this process. Generally, an understanding of the notion of visualization in this context could range from the film itself as a form of visualization, to the many ways one can visualize a film or parts of it, through to using visualization as an explorative tool on both macro and micro levels to prove certain hypotheses, and, finally, to provide aids to viewing and searching entire cultural heritage collections. In this chapter, I will tackle all these points but will focus primarily on the use of visualization as an explorative tool.

Films are usually analyzed according to their visual properties on the one hand or their structured sequentiality on the other. My own specific research interest lies in the application of "visualization without reduction," a concept developed by Manovich, to film works, which means essentially a radical departure from the traditional visualizations using data derived from transcriptions. However, as mentioned previously, the formal analysis of filmic structures has a long history in Slavic Studies, namely in the work of the Russian Formalists. In recent times, with enhanced computational power, it has become possible to utilize the full image of a film rather than just statistical data for analysis aided by visualization. However, as Johanna Drucker (2014) insistently reminds us, visualizations, like ontologies, are always interpretations, as data does not have an inherent visual form which merely gives rise to graphic expression. The type of visualization chosen is therefore of crucial importance, and there are of course ample possibilities for a more statistical or abstract approach to film visualization. For film studies, we can develop a whole other set of possible questions if we also take the temporal aspect into account.

I will begin with an example with which many readers may already be familiar. Lev Manovich and his team created a visualization of all the covers of *Time Magazine* issued between 1923 and 2009. The visualization can be viewed online.[12] How can we create something similar using moving images? Here one needs to take movement out of the equation and create an image sequence consisting of one image for every frame of the film and store it in a folder.[13] In this way, we are creating a kind of visual fingerprint of the entire film or corpus of films we are interested in, and thus we can gain at least superficial visual insights merely from looking at the montage of images. As we know from film studies, there are more sensible ways to break film into meaningful temporal units: frames, shots, scenes, reels and even entire films. These units, with the exception of scenes, can be easily deduced from the film material itself without too much human interpretation. In order to illustrate

12 Manovich and his team explain how it works in this video, which maybe serves as the best introduction to the topic: https://www.youtube.com/watch?v=d_OceOpCmf8
13 The issue of archival frame rates is complicated in itself and beyond the scope of this article. Important to know is that analog film prints from the silent era were not projected at 24fps but usually between 16fps and 24fps, the precise projection speed(s) depending on various factors.

these preliminary notions, I would like to discuss my own collaboration with Lev Manovich on the films of Russian avant-garde director Dziga Vertov.

The project "Digital Formalism" ran from 2007 to 2010 and was a joint endeavor between three institutions coming from three very different disciplines: the Department of Theater, Film and Media Studies at the University of Vienna, the Technical University of Vienna and the Austrian Film Museum. The discipline-specific methods, goals and dissemination traditions of both film and media scholars and computer scientists remained a constant challenge for each of the partners. In the project, eight films by Dziga Vertov were manually annotated using ANVIL by the author, and the data gained served as the Ground Truth for subsequent computer aided analysis. Our basis were the temporal units described above, which we calculated by subtracting the value of the first exported frame of the shot from the value of the last frame. Shots were then tagged with different parameters. The basic aim was to gain insight into the highly formalized artistic work of the director by applying quantitative and formal analysis as well as close readings, and to correlate the data with surviving original documents from the so called Dziga Vertov Collection held at the Austrian Film Museum. [14] However, much of what I will present in this paper has been the result of my own collaboration with Manovich that was carried out after the project had ended, later becoming part of my book *Kollision der Kader* (2016), in which the entire process is described in detail and the visualizations used for my own interpretations and analyzes. Many of the visualizations can also be viewed and explored online. [15]

Movement as one of the key characteristics of cinematographic works is not easy to depict in a static image, and therefore criticism of visualizations for forgoing this vital dimension is valid. However, there is a long tradition in the history of film, going back to Eadweard Muybridge, of breaking down the continuum of movement into discrete images for study. However, there are ways to still retain a sense of time or movement. In the first visualization presented here (see Figure 1), we see a montage of the first frames of every shot in Dziga Vertov's fourth feature-length film "The Eleventh Year" (Dziga Vertov, 1928). This montage was created using the free software ImageJ, [16] by using a Macro from a manually-produced shot breakdown provided in a separate Excel file. The film begins at the top left of the image and ends bottom right. Not only do we see the temporal (chronological) structure of the film, but, depending on the zoom factor, we also get a good impression of image composition and light/dark distribution in the black and white film.

14 For more information see:
https://www.filmmuseum.at/en/collections/special_collections/dziga_vertov_collection. The frame grabs from Vertov's films featured in this paper are taken come from the Austrian Film Museum's collection.
15 For more information see: https://www.flickr.com/photos/culturevis/albums/72157622608431194
16 For more information see: https://imagej.nih.gov/ij/

Figure 1: One frame per shot from "The Eleventh Year" (Dziga Vertov, 1928) (Lev Manovich/Software Studies Initiative). The film begins on the top left of the image and ends on the bottom right.

The same method can be used for both macrostructures (e.g. the whole film) as well as microstructures (e.g. parts of the film). In order to choose relevant sequences for a particular visualization, it is of course helpful, if not necessary, to know the material well. For me, this process has always been one of mutual influence: while performing a close reading of a film something in its form or content would trigger research interests, whereas a previously-made visualization of the entire film might urge me to look closer at particular sequences. This is especially true if we are dealing with filmmakers who, for example, employ a highly elaborated montage technique, such as the Soviet avant-garde directors.

In the following two examples, this time taken from Vertov's fifth and most famous film "Man With a Movie Camera" (Dziga Vertov, 1929) and again from "The Eleventh Year," I have chosen sequences which carry specific messages that are represented formally. It is beyond the scope of this paper to explain in detail how Vertov used slow motion and how the arrangement of the slow motion shots in his films relate to his "Kinoglaz" theory. My aim here is rather to demonstrate that a relatively simple visual representation of one sequence can aid the navigation through a very complex structure. For my purposes, I have chosen to combine a diagram of the shot lengths and a montage visualization (see Figure 2). The sequence shows various people engaging in amateur sports after work while others watch. The shots of the sportsmen and women are captured in slow motion and alternate with the spectators' faces viewed in close up. In the top diagram, each bar represents a single shot; the longer the bar, the longer the shot. Shots in slow motion are highlighted in red to demonstrate visually the way Vertov structures his sequence. In the bottom image, the same sequence is visualized as a montage of the first frame of each shot so as to also be able to see the film's content.
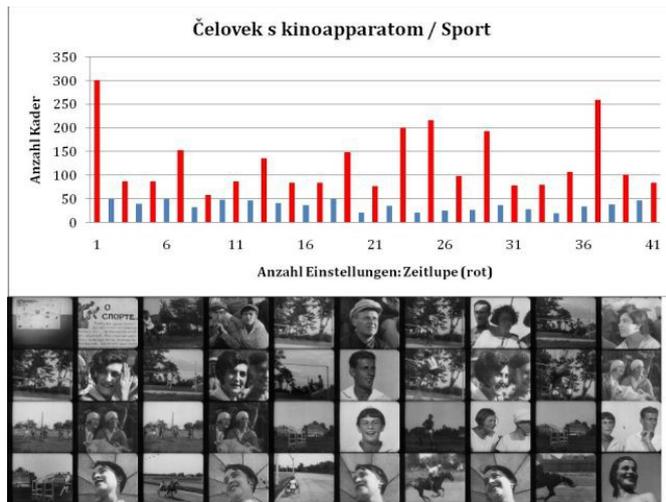
Figure 2: Slow motion sequence in "Man With a Movie Camera" (Dziga Vertov, 1929). One frame per shot, beginning on the top left and ending on the bottom right. Software: ImageJ.[17]

Close-ups of faces have a special significance for Vertov, and he often arranged or filmed them from unusual angles, a method stemming partially from Vertov's constructivist influences, namely his friendship with Aleksandr Rodčenko. The important thing for us to know here is that close-ups of faces in "The Eleventh Year" (see Figure 3) are clustered together in only a few sequences, a formal method Vertov has not used in other films. Manovich (2012) has presented us with some stimulating observations, while I (2016) have tried to locate Vertov's use of faces more comprehensively within the political changes which occurred in Soviet Russia in the 1920s and 1930s. It should be mentioned that Vertov very often used stock footage or re-used his own footage, while at other times he had a camera-team (including, for example, his brother Michail Kaufman) shoot material for him to use. Following traces like these, which are linked both to production as well as to re-use, proved to be a worthwhile film historical investigation for computer scientist Maia Zaharieva (2010) and another potential area of application for digital tools (for example, on film restoration projects, for which different sources very often have to be compared).

Figure 3: All close-ups of faces that appear in "The Eleventh Year" (Dziga Vertov, 1928). One frame per shot, beginning on the top left and ending on the bottom right. Software: ImageJ.

17 If not otherwise stated, the visualizations were created by the author.

The next two images follow the same idea and visualize the editing structure and image composition of a trailer for the silent German crime thriller "Spione" by Fritz Lang (1928). The first visualization (see Figure 4) features every hundredth frame from the image sequence, arranged from top left to bottom right. The montage of the images here conveys an impression of the shot lengths where the number of images depicted corresponds to the length of the shot. In the second visualization (see Figure 5) we focus rather on image composition within shots. The first and the last frame of each shot are here arranged in sequence from top left to bottom right following the chronology of the film. Just by looking at this visualization we start raising questions: Which actions occur within a particular shot and how dynamic are they? Does the composition or the depiction of the leading actors tell us something about the genre or the particular application of the film (trailers as advertising tools)?

Figure 4: Montage of every hundredth frame from a trailer for "Spione" (Fritz Lang, 1928), beginning on the top left and the ending on the bottom right. Software: ImageJ. Source material provided by the Austrian Film Museum.

Figure 5: Montage of the first and last frame of each shot from a trailer for "Spione" (Fritz Lang, 1928), beginning on the top left and the ending on the bottom right. Software: ImageJ. Source material provided by the Austrian Film Museum.

If we are interested in a film's visual properties rather than its chronology, images can be sorted easily according to their degree of brightness. From the following visualization (see Figure 6), again taken from Vertov's "The Eleventh Year," we can learn how many bright and dark images created by the director were used. As Manovich has observed (2013), the opposition between large proportions of

very dark and very light shots can be regarded as specific to "The Eleventh Year": "The former are outside shots, with the sky occupying the larger part of a shot. [...] The dark shots represent industrialization, showing people operating machinery and steelmaking." Such visualizations make comparison between different films easily possible, as Manovich (ibid.) shows: "In contrast, A Man with a Movie Camera takes place in a city, with time covering a whole day from morning to evening. Accordingly, the distribution of its shots is more even, with every gray tone being represented equally."
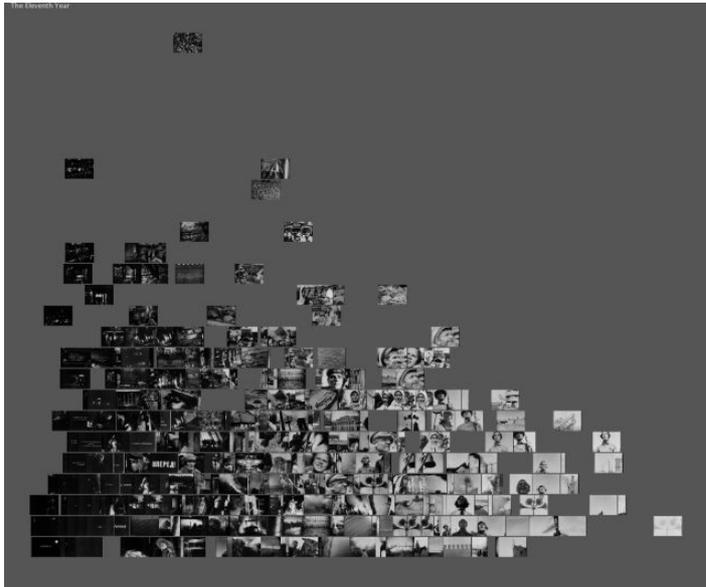

Figure 6: Arrangement of the second frame of every shot from "The Eleventh Year" (Dziga Vertov, 1928) according to visual properties (Lev Manovich/Software Studies Initiative). The x-axis represents the mean (average) gray scale value and the y-axis the number of shapes present.

The research questions posed can be more or less related to film historical topics. In my next two examples, I have tried to visualize the color patterns as featured in two silent films from the late 1910s and the 1920s. Both of these films were tinted; a common procedure at that time, which could be individually tailored to local tastes. The advantage here for those scholars interested in early film color is to be able to see the total number of colors used and the dispersal of the different colors throughout the film in just one single image. In this way, it would be quite feasible to plot the changes in coloring techniques over a specific period of time (say, from the 1920s to the late 1920s). To illustrate this, I have prepared visualizations of a film from 1919 (see Figure 7) and a film from 1926 (see Figure 8).

Figure 7: Montage of every hundredth frame from "Twist Olivér" (Márton Garas, 1919), beginning on the top left and ending on the bottom right. Software: ImageJ. Source material provided by the Hungarian National Digital Archive and Film Institute (MaNDA).
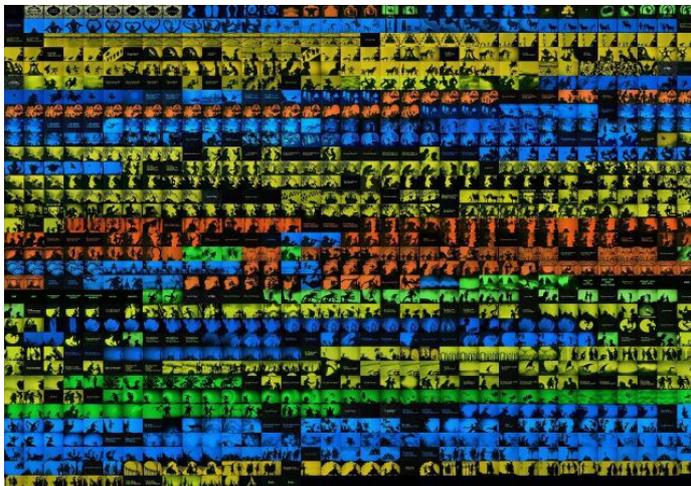

Figure 8: Montage of every hundredth frame from "Die Abenteuer des Prinzen Achmed" (Lotte Reiniger, 1926), beginning of the film is top left and the end is at the bottom right. Software: ImageJ. Source material provided by the Deutsches Filminstitut – DIF / Milestone.

The movement of objects or the camera in one shot is, however, crucial for the understanding of Vertov's oeuvre and cannot be ignored completely. As has already been stated, it is beyond the scope of this article to explain Vertov's theory of film in detail. However, at this point it is worth mentioning that Vertov based his theory on what he called "intervals" (Vertov 2008 [1928], 161), which basically refers to the visual movement between shots. I (2016) have tried to link Vertov's theory to his work by examining the correlations between the different shots in his films according to their formal properties (e.g. composition, camera perspective, movement within the shot, light/dark values and the shooting speed). What Vertov is referring to is essentially a kind of visual rhythm, where the "interval" becomes a collective term for all its different variants. How can we integrate into the static visualizations an impression of the duration of shots in relation to their neighboring shots? How can we compare movement within different shots and present the results visually? How can we correlate shot lengths with movement or images and motifs with movement?

I will start by discussing the formal temporal units which originate from the nature of film production and exhibition during the time in which Vertov made his films. At that time, films were split

into several reels, each measuring approximately 300 meters (or ca. 12 min at a projection speed of 18fps). For his film "Man with a Movie Camera," Vertov had intentionally added clear markers for the start and end of each reel following a distinct pattern, as Yuri Tsivian and myself have investigated (Heftberger 2009). When analyzing his films, it proved necessary to take these reel divisions into account and their relation not only to the structure of the film as a whole but also to the individual shots. "The Eleventh Year" consists of five reels, three of which are depicted in the following visualization (see Figure 9) from bottom to top, meaning reel one is on the bottom, reel two in the middle and reel three at the top. In order to be able to see the details more clearly here, I have isolated only the beginning part of each reel. Underneath the first frame of the shot a white bar representing shot length: the longer the bar, the longer the shot. The shots are depicted in chronological order moving from left to right. Immediately apparent are the long shots at the beginning of the film (bottom left), due to the explanatory intertitles and credits. The subsequent intertitles composed of only a single word each become part of the temporal rhythm, alternating with longer shots. Zooming in on the formally striking parts of the visualization in this way allows us to develop theories as to the chosen structure.
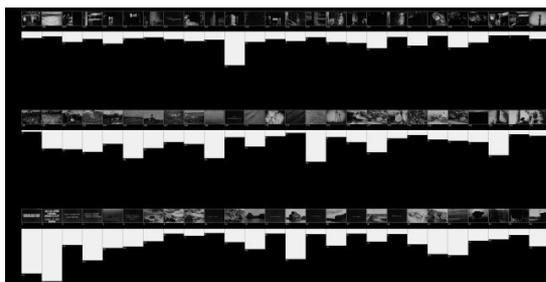

Figure 9: Shot lengths in three 300 meter reels of "The Eleventh Year" (Dziga Vertov, 1928) (Lev Manovich/Software Studies Initiative). The bars underneath the frames represent the shot length.

Measuring the movement within shots is difficult, and algorithms rarely prove reliable, especially when it comes to historical prints that suffer from (at times, extreme) amounts of wear and tear. Cutting (2011, 571) and Manovich (2013) have found useful ways to attain approximate results which can be used for explorative purposes. In my next example (see Figure 10), bars are again visible underneath the frames, but this time they represent the movement within the shot: the longer the bar, the more dynamic the shot.
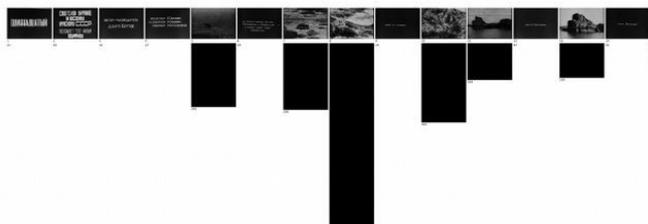
Figure 10: Movement within the opening shots of "The Eleventh Year" (Dziga Vertov, 1928) (Lev Manovich/Software Studies Initiative). The longer the bar, the more vivid the action within the shot. The movement was calculated automatically by Manovich.

An alternative approach would be to tag the movement in a film manually following a self-defined system, as I did on an experimental level with "The Eleventh Year". Here I assigned each shot a rather generic category such as: no motion, slow motion camera, slow motion naturally, normal motion naturally, fast motion naturally, fast motion camera and irrelevant. The following example (see Figure 11) depicts the end of "The Eleventh Year," for which the manual tagging has been combined with yet another form of visualized movement. The film is depicted chronologically from left to right, and the higher the image is situated on the vertical axis, the more movement occurs within the shot. Here we clearly observe that the film becomes more animated the closer it comes to the end. This montage pattern, which Vertov consistently applied throughout his films, can in my view be depicted rather convincingly in visualizations such as this as well as the previous one (see Figure 10).

Figure 11: Movement within the final shots of "The Eleventh Year" (Dziga Vertov, 1928) (Lev Manovich/Software Studies Initiative). The higher up the image, the more dynamic the movement. The degree of movement was tagged manually on a scale from one to seven.

The next and last visualization (see Figure 12) combines the results of three different visualization experiments. The goal was to visualize the so-called "episodes" (Vertov's own term to describe the semantic as well as the formal units in his films) in terms of shot length, movement within shots and shot composition, all at the same time. In this example from "The Eleventh Year," I have chosen a short episode in which close-ups of faces are repeatedly cross cut. The bottom and middle parts follow a visualization convention already familiar from previous images (see Figure 9 and Figure 11 respectively). The top part, meanwhile, depicts two frames, the first and last of each shot, arranged one on top of the other. Vertov handles shots of faces in a fairly static manner, as the visualization makes apparent. Possible reasons for this have been discussed by Manovich (2013) and myself (2016) in more detail elsewhere. My goal with this experiment was simply to show three possible ways of visualizing one episode in a single image. From the direct comparison, it is possible to deduce those visualizations that can be combined in useful ways, and those which provide an overabundance of potentially confusing information.

Figure 12: Three different ways of visualizing image content and shot length in an excerpt from "The Eleventh Year" (Dziga Vertov, 1928) (Lev Manovich/Software Studies Initiative). The sequence features one frame for every shot arranged from left to right.

To date, my research has focused on fairly small corpora, e.g. the eight films made by Dziga Vertov between 1924 and 1934. However, it is a common fact that one of the major challenges facing researchers as well as curators is the ability to search and browse large corpora and to come up with meaningful results. Here it would be useful to provide some facts about the rapidly growing digital content being produced on a daily basis, which can arguably be called the cultural heritage of our times. Lev Manovich estimates that 300 million photos are shared every day on Facebook, and 80 million photos on Instagram. Many of these photos are highly stylized and thus often referred to as digital art. Jeffrey Schnapp puts these figures into perspective with his statement that, "[e]very two minutes we now take as many photographs as were taken during the entire 19th century."[18] For video the situation is similar, as Luke McKernan, lead curator of moving image at the British Library, points out on his blog: "I estimate that there have been 2.7 billion videos uploaded to YouTube since 2005. 400 hours of video are added to the site every minute". He then compares these figures to film archives, which by his estimate, "haven't managed to collect more than 400 hours of content in years." How can we curate these large corpora? What should be preserved, and how can we ensure it remains accessible (and to whom)? What we are essentially witnessing here is yet another paradigm shift from manual selection carried out by an elite group of experts to a democratic big data model. McKernan aptly describes the current situation facing the cultural heritage institutions thus:

> Vast amounts of this online content is what might be termed trivia: ephemeral videos of skateboarding pets of the kind that would never have been acquired by a film archive, nor even conceived of as a type of film production before the YouTube era. But is it trivia? How are we to judge what a moving image should be? Is the understanding of it as an art medium, of the kind best revered in a cinematheque, now something absurdly narrow? What, intrinsically, is the difference between, say CITIZEN KANE and CHARLIE BIT MY FINGER? Perhaps we should only look at the numbers – unless it is the numbers that are scaring us, and we prefer to cling to old certainties.

While a democratization of collection policies together with re-thinking traditional curatorial self-conceptions are both desirable and necessary, there remains the valid question of meaningful selection. As Schnapp has stated: "How can we find corpora that matter to a given community or within a given cultural domain accessible and usable in a meaningful way?" In order to select material, we need to find it, and ideally we need to be able to locate and access a vast amount of data (if not all

18 The lecture was called "Knowledge Design" and held at the Österreichische Akademie der Wissenschaften in April 2016.

of it) from which to chose: no easy feat. Manovich (2011) claims that this is the fault of current interface designs, since, "[p]opular web interfaces for massive digital media collections such as 'list', 'gallery', 'grid', and 'slide show' do now allow us to see the contents of a whole collection." Since we are not able to view and explore the complete collection's metadata (either online or on site), according to Manovich (ibid.), "without any preconceived expectations or hypotheses, a researcher has to postulate beforehand what the important types of information worth seeking out are." These search strategies have another disadvantage: they do not reveal the context of the objects we are interested in (which subset do they belong to, which are the similar objects etc). How, therefore, can we discover interesting things in massive media collections? How can we browse through them efficiently and effectively without any preconceived notions about what we wish to find? Or, in the case of film archives: How can we link documents or media that were hitherto unknown to us and not linked according to their visual properties or metadata?

Conclusions and Outlook

The formal analysis of a single film does not only tune our eyes to look for specific structures within it, but it also potentially allows for comparisons of several films. It is especially in the field of formal analysis that we can benefit hugely from higher computational power, new digital tools and a methodology that has been developed across different disciplines (e.g. computer sciences, linguistics, visual studies, art history). Investigations into film form can also help to gain greater insight into the historical practices surrounding film production, film aesthetics, film distribution, censorship and not least in the history of a specific film element or print. It would be fascinating in general to be able to combine automatic analysis with more traditional humanist methods (film and TV analysis, close readings, psychoanalysis, discourse analysis etc.).

My overall impression is that computer sciences have already solved many technical issues in terms of automatic analysis, at least where recent film production with crisp images and clear sound are concerned, but the real problem lies in its practical application for film scholars; something which is still missing. Another obstacle I observe is a large gap in knowledge transfer from the computer sciences, so that film and media studies have a better idea of which is already possible in image analysis like shot recognition, image recognition, etc. The differences between scholarly practices present another, related problem for interdisciplinary research proposals: that which humanities' scholars would like to quantify or analyze is often either too simple for computer scientists or far too complex. Often knowledge transfer does not occur as fast as it could, as even in interdisciplinary teams algorithms remain firmly within the technical realm and cannot be used by humanists so easily. Sometimes other computer scientists are not able to adapt the existing methods for their own needs easily (algorithm black box). Without delving too deeply into the topic of Open Data and Open Science, I would still strongly argue that it is now the time to build research infrastructures for sharing data within film studies.

Traditional qualitative methods can help us to define what we are looking for and whether or not it is valuable in the given context, especially when the goals go beyond merely recognizing

day/night, trees and houses, etc. This is something computer scientists and humanists need to work on together, jointly tackling issues such as: How can we create and visualize meaningful data which will then aid discourse analysis, cultural studies or, in this case, Russian Formalism? When should we apply qualitative analysis and when not? Or should we opt for a more explorative approach, let computer scientists do a basic search and pre-rank the results for humanists to evaluate? As has been pointed out by Mark Williams (2016, 344), the digital humanities, "regularly features an iterative dialectic between the traditions of 'close reading' in the arts and humanities versus the goals and practices of 'distant reading' crucial to computational approaches to vast corpora of media texts under analysis. Recognizing these sites of potential dissonance will continue to be fundamental to progress in the emerging interdisciplinary space that is DH."

Data maps are still as popular as ever in the humanities (Moretti 2005) and seem an obvious choice when it comes to film studies, whether tracking physical film prints or the protagonist in the course of the on-screen narrative. We can either create the filmic geography or try to map the narrative, etc. These representations of space often go hand in hand with investigations into sociopolitical tendencies, as scholars like Oksana Bulgakowa (2003) or Emma Widdis (2003) have done for Russian cinema of the 1920s and 1930s. Laura Horak mentions the influential work by Deb Verhoeven and her team (Verhoeven 2009, 79), who have stated that mapping, "offers most when it raises new questions about spatial and temporal connectivity, rather than promising closure on the question of what was going on in the past."[19] However, Johanna Drucker insists (2014, 77) that, "the greater intellectual challenge is to create spatial representations without referencing a pre-existing ground."

Film, of course, is a time based media form and this aspect of its character should be included in some way, for example in time series' or narrative graphics of space and time. Representation of time in art is a complex matter, as becomes instantly apparent if we just start to consider which form of time we actually want to represent: the time of the film's creation, the time depicted in it, the running time, etc. Another interesting idea would be to push new ways of defining the curation of film programs. Is it imaginable to develop algorithms as curatorial tools that would automatically produce a list of options from certain pre-defined parameters (keywords, directors, genre, time period, or even colors or other visual qualities)? While this might be something that could be used by film archives for their online presentations (along the lines of a "video of the month" or "you might also like this" feature), this form of curation without a curator will probably be met with reservation by other quarters (e.g. cinema programmers). However, it would potentially break down personality-driven choices and help bring material to the foreground, which is otherwise rarely if ever shown.

All the enthusiasm for data visualization should not mask the fact that also much criticism and skepticism has been expressed. Johanna Drucker (2014) claims, for example, that the introduction of representational practices from the natural sciences also potentially introduce scientist and positivist notions to the historian's practice anew, and she calls for a more pronounced input from the humanities: "They need a way to graph and chart temporality in an approach that suits basic

19 In her article "Using Digital Maps to Investigate Cinema History", Laura Horak (2016) provides an impressive list of resources for mapping software.

principles of interpretative knowledge." Drucker (2014, 54) reminds us that all graphical schemata are built on the single principle of defining classes of entities and of relations, especially in the realm of the semantic web. For a humanistic approach, these have to be defined as rhetorical arguments produced as a result of making, a poetics of graphical form, not in the reductive or abstract logic of Boolean algebra. In other words, Drucker arrives at a fully humanistic system for visualizing interpretation. If we take her criticism seriously, we need to think more thoroughly about the humanities' input when it comes to the design of interfaces and websites, drawing on longstanding experience in knowledge design and expression of complex semantic issues. Drucker foregrounds the interpretative subjective aspect of the humanities, so why not develop individual search entries, explore narratives or artistic approaches?

On the other hand, as distinguished visualization specialist Moritz Stefaner (Bihanic 2013) likes to point out, his visualizations always tell more than one story, and, "[t]he trick is to not present them all simultaneously or with the same priority, but deliberately establish a hierarchy and sequence of perception events." Storytelling is used by Jeffrey Schnapp at Harvard's metaLab, where he and his team try to develop meaningful tools for the visualization of cultural objects. Schnapp (2014) reminds us that visualizations which incorporate faces and human-centric scenes have been found to be more powerful, particularly when embedded within stories. Using the term storied collections, he proposes to build an infrastructure and tools which first and foremost support the humanistic culture of critical engagement with data.[20] Along these lines, he (ibid., 12) poses the questions, "how do we weave outputs crafted with such tools into forms of argument and narration that signify culturally, that tell stories of consequence, that support or even replicate the magic and enchantment of traditional forms of storytelling? How do we effectively embed human faces into trees, networks and matrices? And what sort of distinctive new types of stories do collections want to tell that they have been unable to tell with prior toolkits?"

In order to have more material at our disposal for the kind of studies and publications described above, and to seriously engage with big corpora of (high quality) moving images, there are a number of obstacles which must first be overcome. Firstly, archives require continuous financial support if they are to be able to preserve and digitize their collections. They also need support from the academic community in their fund raising efforts and to make their work meaningful and visible. In the end, only through collaboration between cultural heritage institutions and research institutions combined with public participation can the film heritage be made available for scholars and the general public alike. Secondly, we need to establish research infrastructures where archivists and curators can participate rather than merely being viewed as (unwilling) content providers. For are they, in the end, not the experts when it comes to understanding the medium, whether it is analog or digital?

20 Schnapp describes his project "Curarium", designed with the intention of fostering collaboration between a diverse community of users: https://curarium.com/. One recent project for presenting museum objects and their metadata is in a convincing way is "Lightbox": https://vimeo.com/158652479. Another example comes from the New York Public Library, which presents digital objects in innovative, interactive ways: http://publicdomain.nypl.org/pd-visualization/.

References

[Balázs 1982 (1925)] Balázs, B. *Schriften zum Film*. Henschelverlag Kunst und Gesellschaft, Berlin (1982 [1925]).

[Bihanic 2013] Bihanic, D. "An Interview with Moritz Stefaner." May 2013. Retrieved from: http://truth-and-beauty.net/appearances/in-the-media/new-challenges-for-data-design.

[Birett 1988] Birett, H. "Alte Filme: Filmalter und Filmstil. Statistische Analyse von Stummfilmen." In: E. Ledig (ed), Der Stummfilm. Konstruktion und Rekonstruktion. diskurs Film, München (1988), pp. 69-88.

[Bordwell 1985] Bordwell, D., Staiger, J. and Thompson, K. *The Classical Hollywood Cinema. Film Style & Mode of Production to 1960*. Columbia University Press, New York (1985).

[Bulgakowa 2003] Bulgakowa, O. "Spatial Figures in Soviet Cinema of the 1930s". In: E. Dobrenko and E. Naiman (eds), The Landscape of Stalinism. The Art and Ideology of Soviet Space, Seattle, London (2003), pp. 51-76.

[Burghardt 2016] Burghardt, M., Kao, M. and Wolff, C. "Beyond Shot Lengths – Using Language Data and Color Information as Additional Parameters for Quantitative Movie Analysis", Conference Abstracts. Jagiellonian University & Pedagogical University, Kraków (2016): pp. 753-755.

[Cutting 2011] Cutting, J., Brunick, K., DeLong, J., Iricinschi, C. and Candan, A. "Quicker, faster, darker: Changes in Hollywood film over 75 years", i-Perception 2 (2011): pp. 569-576.

[Cutting 2013] Cutting, J., Iricinschi, C. and Brunick, K. "Mapping Narrative Space in Hollywood Film", Projections 7/2, Winter (2013): pp. 64-91.

[Drucker 2014] Drucker, J. *Graphesis. Visual Forms of Knowledge Production*. Harvard University Press, Cambridge (2014).

[Ėjchenbaum 2005 (1927)] Ėjchenbaum, B. "Probleme der Filmstilistik." In: W. Beilenhoff (ed), Poėtika Kino. Theorie und Praxis im russischen Formalismus, Frankfurt am Main (2005 [1927]), pp. 20-55.

[European Commission 2014] European Commission. "Film Heritage in the EU. Report on the Implementation of the European Parliament and Council Recommendation on Film Heritage 2012-2013." 1 October 2014, pp. 1-39. Retrieved from: https://ec.europa.eu/digital-single-market/en/news/european-commissions-report-film-heritage.

[Heftberger 2009] Heftberger, A., Tsivian, Y. and Lepore, M. "*Man with a Movie Camera* (SU 1929) under the Lens of Cinemetrics". In: K. Gruber, B. Wurm, V. Kropf (eds), Digital Formalism. Die kalkulierten Bilder des Dziga Vertov. Maske und Kothurn 55/3 (2009): pp. 31-50.

[Hanley, Heftberger 2012] Hanley, O. and Heftberger, A. "Scholarly Archivists/Archival Scholars: Rethinking the traditional models", The Velvet Light Trap 70 (2012): pp. 64-65.

[Heftberger 2014] Heftberger, A. "Film archives and digital humanities – an impossible match? New job descriptions and the challenges of the digital era", MedieKultur – Journal of media and communication research 30, 57 (2014): pp. 135-153.

[Heftberger 2016] Heftberger, A. *Kollision der Kader. Dziga Vertovs Filme, die Visualisierung ihrer Strukturen und die Digital Humanities*. edition text+kritik, München (2016).

[Horak 2016] Horak, L. "Using Digital Maps to Investigate Cinema History." In: E. Hoyt and C. Acland (eds), The Arclight Guidebook to Media History and the Digital Humanities, Sussex (2016), pp. 65-102.

[Hoyt 2014] Hoyt. E. "How to Topic Model a Fan Magazine." 17 November 2014. Retrieved from: http://projectarclight.org/arguments/how-to-topic-model-a-fan-magazine/.

[Hoyt 2016] Hoyt, E., Hughes, K. and Acland, C. "A Guide to the Arclight Guidebook." In: E. Hoyt and C. Acland (eds), The Arclight Guidebook to Media History and the Digital Humanities, Sussex (2016), pp. 1-29.

[Jacobs 2014] Jacobs, L. *Film Rhythm after Sound. Technology, Music, and Performance*. University of California Press, Oakland (2014).

[Kaes 2016] Kaes, A., Baer, N. and Cowan, M. *The Promise of Cinema. German Film Theory 1907 - 1933*. University of California Press, Oakland (2016).

[Kazanskij 2005 (1927)] Kazanskij, B. "Die Natur des Films." In: W. Beilenhoff (ed), Poėtika Kino. Theorie und Praxis im russischen Formalismus, Frankfurt am Main (2005 [1927]), pp. 86-129.

[Kudenov 2016] Kudenov, P. "Interviewing Mark C. Marino: HASTAC, Digital Humanities, and Critical Code Studies." 11 October 2016. Retrieved from:
https://www.hastac.org/blogs/peter-kudenov/2016/10/11/interviewing-mark-c-marino-hastac-digital-humanities-and-critical.

[Manovich 2011] Manovich, L. "Against Search." Retrieved from: http://lab.softwarestudies.com/2011/07/against-search.html.

[Manovich 2012] Manovich, L. "Trending." In: M. Gold (ed), Debates in the Digital Humanities, Minneapolis, London (2012), pp. 460-475.

[Manovich 2013] Manovich, L. "Visualizing Vertov." 2013. Retrieved from: http://softwarestudies.com/cultural_analytics/Manovich.Visualizing_Vertov.2013.pdf.

[Moretti 2005] Moretti, F. *Graphs, Maps, and Trees. Abstract Models for a Literary History*. Verso, London (2005).

[O'Brien 2005] O'Brien, C. *Cinema's Conversion to Sound: Technology and Film Style in France and the U. S*. Indiana University Press, Bloomington, Indiana (2005).

[Olesen n.Y.] Olesen, C. "Formalising Digital Formalism: An Interview with Adelheid Heftberger and Matthias Zeppelzauer about the Vienna Vertov Project." In: M. de Rosa and L. Fales (eds), Shifting Layers. New Perspectives in Media Archaeology Across Digital Media and Audiovisual Arts, Milan (forthcoming 2016).

[Redfern 2013] Redfern, N. "Film Studies and Statistical Literacy", Media Education Research Journal 4(1) (2013): pp. 58-73.

[Salt 1992] Salt, B. *Film Style and Technology: History and Analysis*. Starword, London (1992).

[Schnapp, Presner 2009] Schnapp, J. and Presner, T. "Digital Humanities Manifesto 2.0." 2009. Retrieved from: http://humanitiesblast.com/manifesto/Manifesto_V2.pdf.

[Schnapp 2014] Schnapp, J. *Knowledge Design. Incubating new knowledge forms / genres / spaces in the laboratory of the digital humanities*. Herrenhauser Lectures: Hannover (2014).

[Šklovskij 1984 (1925)] Šklovskij, V. "Literatur ohne Sujet." In: V. Šklovskij, Theorie der Prosa, Frankfurt am Main (1984 [1925]).

[Timoschenko 1928] Timoschenko, S. "Filmkunst und Filmschnitt." In: W. Pudowkin, Filmregie und Filmmanuskript, Berlin (1928).

[Verhoeven 2009] Verhoeven, D., Bowles, K. and Arrowsmith, C. "Mapping the Movies: Reflections on the Use of Geospatial Technologies for Historical Cinema Audience Research." In: M. Ross, M.

Grauer, and B. Freisleben (eds), Digital Tools in Media Studies: Analysis and Research: An Overview, Bielefeld and New Brunswick (2009), pp. 69-81.

[Vertov 2008 (1928)] Vertov, D. "Čto takoe Kino-Glaz." In: D. Kružkova (ed), Dziga Vertov iz nasledija. Tom vtoroj. Stati i vystuplenija, Moscow (2008 [1928]), pp. 157-164.

[Walkowski 2016] Walkowski, N.-O. "The Landscape of Digital Annotation and Its Meaning." Conference Abstract on Language Technologies & Digital Humanities, Ljubljana (2016): pp. 6-11.

[Weingart 2016] Weingart, S. "Lessons from Digital History's Antecedents." 30 October 2016. Retrieved from: http://scottbot.net/lessons-from-digital-historys-antecedents/.

[Widdis 2003] Widdis, E. *Visions of a New Land. Soviet Film from the Revolution to the Second World War*. Yale University Press, New Haven and London (2003).

[Williams 2016] Williams, M. "Networking Moving Image History: Archives, Scholars, and the Media Ecology Project." In: E. Hoyt and C. Acland (eds), The Arclight Guidebook to Media History and the Digital Humanities, Sussex (2016), pp. 335-346.

[Zaharieva, Breiteneder 2010] Zaharieva, M., and Breiteneder, C. "Archive Film Comparison", International Journal of Multimedia Data Engineering and Management 3/1 (2010): pp. 41-56.