Tim van der Heijden, Taylor Arnold, and Lauren Tilton

Distant Viewing the Amateur Film Platform

Introduction

This chapter explores the Amateur Film Platform (2014–2023), an online platform that hosted a unique collection of more than eight thousand films and videos made by Dutch amateur filmmakers in the twentieth century.[1] The Amateur Film Platform was initiated by the Netherlands Institute for Sound & Vision (NISV) in collaboration with various regional audiovisual archives in the Netherlands, including the City Archive Rotterdam and the Groningen Audio Visual Archive (GAVA). After the platform's launch in 2014, the Frisian Film Archive, Drenthe Audio Visual Archive (DAVA), Limburgs Museum, and Eye Filmmuseum also presented parts of their amateur collections on the platform, which made the platform diverse, geographically, institutionally, and historically – featuring amateur footage from the early 1900s until the 2010s from various technological carriers. The Amateur Film Platform predominantly presented digitized film-based collections, including 16mm, 9.5mm, 8mm, and Super 8 films, but digitized electronic videos and born-digital materials were also included, such as VHS, MiniDV, and web videos.[2] Moreover, via Community Uploads, users could upload their own (digitized) films and videos directly to the platform's portal, which made the platform function as a "living archive" of both historical and contemporary amateur media productions (Figure 1).[3]

Prior to the Amateur Film Platform's discontinuation in 2023, we have explored the platform's collection by utilizing the approach of Distant Viewing, which as a

1 See https://www.amateurfilmplatform.nl/, accessed September 29, 2022. The platform was online from 2014 until 2023. For more details about the platform's discontinuation and how to retrieve some of its collections, see https://www.beeldengeluid.nl/kennis/projecten/amateurfilmplatform, accessed March 22, 2024. Captures of the platform throughout the years can be found via the Internet Archive's Wayback Machine: https://web.archive.org/web/20230315000000*/https://www.amateurfilmplatform.nl/, accessed March 22, 2024.

2 For more details on these different amateur media and technological formats, see Tim van der Heijden, "Technologies of Memory: Amateur Storage Media and Home Movie Practices in the Longue Durée," *Le Temps Des Médias* 39, no. 2 (2022): 141–159, https://doi.org/10.3917/tdm.039.0141; Leo Enticknap, *Moving Image Technology: From Zoetrope to Digital* (London, New York: Wallflower, 2005). See also: Tim van der Heijden and Valentine Kuypers, "Life in Motion: A History of Amateur Film," *Europeana*, April 12, 2023, accessed August 8, 2024, https://www.europeana.eu/en/exhibitions/life-in-motion/.

3 Cf. Susan Aasman, "Finding Traces in YouTube's Living Archive: Exploring Informal Archival Practices," *TMG Journal for Media History* 22, no. 1 (2019): 35–55, https://doi.org/10.18146/tmg.435.

https://doi.org/10.1515/9783111082486-012

methodological framework applies computer vision methods to the computational analysis of large collections of audiovisual materials.[4] Distant Viewing enables us to work with large corpora of audiovisual materials to extract, aggregate, and visualize certain features or "semantic elements of visual materials," such as color, shot length, object detection, and camera movement.[5] It makes use of the Distant Viewing Toolkit (DVT), which is designed to facilitate the computational analysis of visual culture employing the latest machine learning and computer vision techniques.[6] The software, published as an open-source library, consists of several Python packages of computer vision algorithms, which can be used to navigate audiovisual collections and analyze their features at scale.[7] What makes Distant Viewing distinct from other digital approaches to visual culture is that it makes "explicit the interpretative nature of extracting semantic metadata from images," reflecting on how the computational viewing of audiovisual materials is informed by a certain code system.[8]

By applying the Distant Viewing approach and toolkit to the Amateur Film Platform, we aim to explore how it can provide new insights into this relatively large collection of amateur films and videos from the Netherlands. More specifically, we aim to investigate whether any formal, stylistic, and aesthetic patterns or changes over time could be found in the Amateur Film Platform collection, based on the data sample that we used for the analysis and, if so, how these patterns could contribute to a broader understanding of the history of amateur film as a cultural practice.[9]

4 Taylor Arnold and Lauren Tilton, *Distant Viewing: Computational Exploration of Digital Images* (Cambridge, MA: MIT Press, 2023); Taylor Arnold and Lauren Tilton, "Distant Viewing: Analyzing Large Visual Corpora," *Digital Scholarship in the Humanities* 34, issue supplement 1 (2019): i3–i16, https://doi.org/10.1093/llc/fqz013.

5 Arnold and Tilton, "Distant Viewing," i6.

6 Distant Viewing Toolkit, https://github.com/distant-viewing/dvt, accessed March 22, 2024.

7 For the software white paper, see Taylor Arnold and Lauren Tilton, "Distant Viewing Toolkit: A Python Package for the Analysis of Visual Culture," *Journal of Open Source Software* 5, no. 45 (2020): 1–6, https://doi.org/10.21105/joss.01800. For a review of the Distant Viewing Toolkit, see Patrick Sui, "Review: Distant Viewing Toolkit," *Reviews in Digital Humanities* 2, no. 4 (2021), https://doi.org/10.21428/3e88f64f.01fedc58.

8 Arnold and Tilton, "Distant Viewing," i4.

9 Van der Heijden conducted a historical discourse analysis and archival research to study the relationship between changing media technologies (film, video, digital media) and home movie practices from a long-term historical perspective. This media historical research informed the hypotheses of the tested relationships in the present study. From a methodological perspective, we were interested in the question of whether the Distant Viewing approach could confirm or provide alternative views on those relationships. See Tim van der Heijden, "Hybrid Histories: Technologies of Memory and the Cultural Dynamics of Home Movies, 1895–2005" (PhD diss., Maastricht University, 2018).

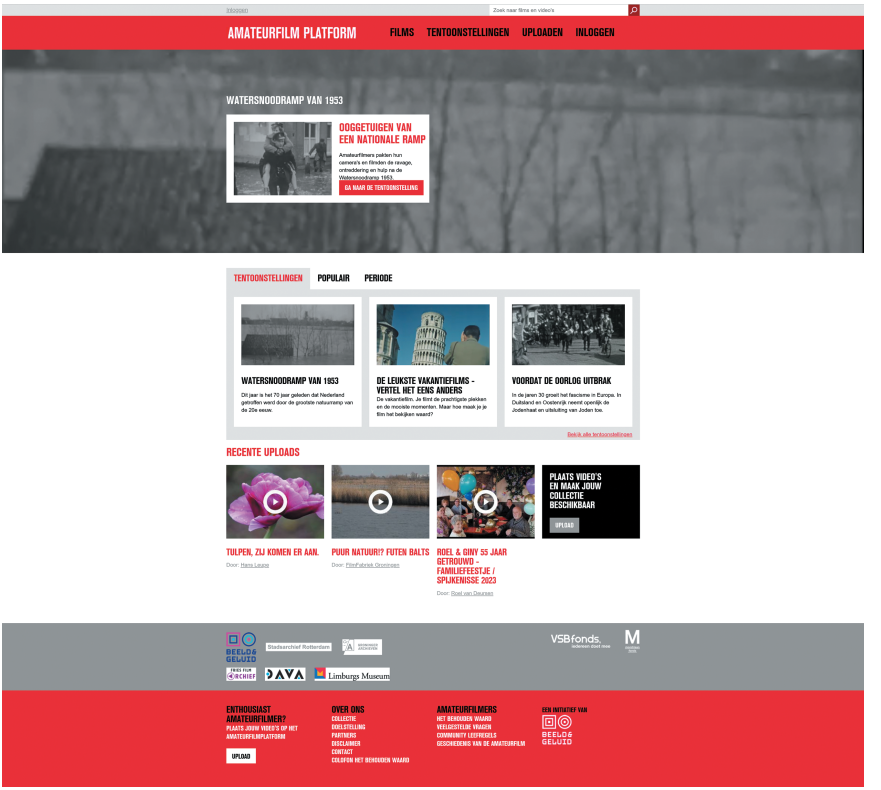

Figure 1: Homepage of the Amateur Film Platform, https://www.amateurfilmplatform.nl/ (accessed April 2, 2023).

Studying Amateur Media Collections at Scale

The Amateur Film Platform presented a large collection of digitized and born-digital amateur media content, produced by Dutch amateur filmmakers over different time periods and by means of various media technologies. The amateur films and videos were digitally searchable as part of a large dataset, which allowed for filtering based on predefined categories and other metadata. What new possibilities did this bring for studying the history of amateur film in the Netherlands? What patterns, relations, and changes over time could we analyze, for instance, in the use of certain topics, genres, narrative, or aesthetic tropes among other categories? In short, what would the potential and challenges be for exploring audiovisual collections of amateur films and videos at scale?

The platform was unique in the sense that it combined amateur collections from various audiovisual archives and presented them online for audiences to navigate and explore in various ways.[10] The user interface facilitated searching the collections by keywords in the search bar at the top, but users could also navigate the platform through various exhibitions that highlighted certain themes and films from the collection, or by means of the films themselves and their metadata. The different types of technological carriers were reflected in the metadata under the category CARRIER, which included multiple film-based formats like 16mm, 9.5mm, 8mm, and Super 8 film as well as video-based formats, such as VHS and MiniDV.[11] Furthermore, the platform allowed searching based on LOCATION (e.g., Amsterdam, Rotterdam, France), TIME PERIOD (from the 1900s until the 2010s), SUBJECT (e.g., amateur recordings, holiday, beaches, families, etc.), FILMMAKER (e.g., Piet Schendstok, Dick Laan, Roel van Deursen), and the archival COLLECTION the film or video items originally belongs to (e.g., NISV, GAVA). All these (sub)categories functioned as entities that could be clicked on to navigate the collection on the platform. To filter the results, categories could be combined in the search menu (e.g., selecting all 8mm films from the 1960s, or all films about the subject families from a specific filmmaker or collection) (Figure 2).

By foregrounding a digital approach to amateur film history, our project relates to other digital projects that involve large historical amateur film collections and archival materials, including projects from the Prelinger Archives,[12] the Bas-

10 Comparable initiatives in northern America include the Prelinger Archives (https://archive.org/details/prelinger, accessed March 22, 2024) and the Amateur Movie Database (an international platform celebrating the world of amateur cinema, which functions as a web-based tool for researching amateur films, filmmakers, and movie clubs from various international archives and collections: https://www.amateurcinema.org/, accessed March 22, 2024). For reflections on the Amateur Movie Database, see Charles Tepperman, "The Amateur Movie Database," *Screen* 61, no. 1 (2020): 124–128, https://doi.org/10.1093/screen/hjaa007; Charles Tepperman, "The Amateur Movie Database: Archives, Publics, Digital Platforms," *The Moving Image: The Journal of the Association of Moving Image Archivists* 17, no. 2 (2017): 106–110, https://doi.org/10.5749/movingimage.17.2.0106.

11 There was also a separate Video subcategory in addition to the VHS and MiniDV subcategories within the CARRIER category, which did not necessarily overlap. Within this article, we therefore use "Video" with a capital V to indicate items falling under this subcategory in the metadata, whereas we use "video" as a more descriptive term for video-based carriers in general.

12 Official website of the Prelinger Archives, an online database founded by archivist and filmmaker Rick Prelinger: http://www.panix.com/~footage/, accessed August 30, 2023. The collection includes more than 17,000 home movies, recorded on 35mm, 16mm, 8mm, Super 8, and other film formats. Currently, 8,596 films from the Prelinger Archives are accessible via the Internet Archive: https://archive.org/details/prelinger, accessed August 30, 2023.

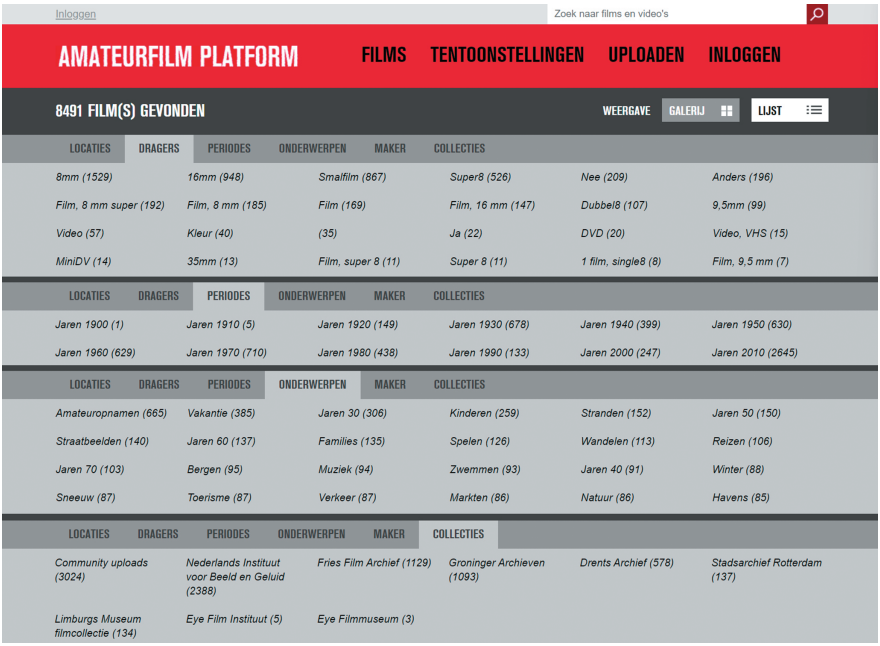

Figure 2: Search menus of the Amateur Film Platform, https://www.amateurfilmplatform.nl/ (accessed April 2, 2023).

que Films Project from the Amateur Movie Database (AMDB),[13] and the Play the City project from the Home Movies Italian Amateur Film Archive in Bologna (Archivio Nazionale del Film di Famiglia), which uses geodatabases to remediate amateur films into on-site installations and exhibitions, allowing visitors to navigate historical maps guided by historical amateur footage.[14] While such projects have

13 The Basque Films Project is a collaboration between Elias Querejeta Zine Eskola (EQZE) and the Amateur Movie Database (AMDB), which develops a detailed map of amateur filmmaking in the Basque Country: https://www.amateurcinema.org/index.php/basquefilms, accessed August 30, 2023.

14 Paolo Simoni, "The Amateur City: Digital Platforms and Tools for Research and Dissemination of Films Representing the Italian Urban Landscape," *The Moving Image: The Journal of the Association of Moving Image Archivists* 17, no. 2 (2017): 111–118, https://doi.org/10.5749/movingimage.17.2.0111. Home movies' ability to function as "time machines" was also explored by artist and digital humanities researcher Ruxandra Lupu in her "Home Movie 4.0" project. See Ruxandra Lupu, "The Home Movie 4.0: (Co)Creative Strategies for a Tacit, Embodied and Affective Reading of the Sicilian Home Movie Archive" (PhD diss., University of Leeds, 2020), accessed August 30, 2023, https://etheses.whiterose.ac.uk/27966/. See also the project website: http://homemoviesicily.com/, accessed August 30, 2023.

been successfully utilizing the affordances and "logics"[15] of databases to abstract and recontextualize amateur film collections and their national or regional histories, they do not make use of digital methods for computationally analyzing the historical films themselves and their audiovisual features or characteristics. This is in line with digital film historical scholarship and approaches in digital humanities research in general, which have predominantly been text-driven and context-oriented rather than focused on the audiovisual materials themselves.[16] Only recently have computer vision methodologies been applied to humanities scholarship, including the analysis of the content and style of audiovisual sources.[17]

While computer vision methodologies have been applied – though in a limited way so far – to archival audiovisual collections,[18] their potential for the study of historical amateur film collections has not yet been fully realized. In general, the use of digital approaches to amateur films and other substandard or non-professional media is a promising yet rather unexplored territory.[19] An explanation for this is that amateur films are generally rather diverse and complex cultural objects that,

15 Lev Manovich was among the first media theorists to write about the logics of databases and the affordances of database-driven narration and storytelling. See Lev Manovich, *The Language of New Media* (Cambridge, MA: MIT Press, 2001).

16 Cf. Charles R. Acland and Eric Hoyt, eds., *The Arclight Guidebook to Media History and the Digital Humanities* (Falmer: REFRAME Books, 2016), accessed March 22, 2024, https://projectarclight.org/book/.

17 Arnold and Tilton, *Distant Viewing*; Lev Manovich, *Cultural Analytics* (Cambridge, MA: MIT Press, 2020); Melvin Wevers and Thomas Smits, "The Visual Digital Turn: Using Neural Networks to Study Historical Images," *Digital Scholarship in the Humanities* 35, no. 1 (2020): 194–207, https://doi.org/10.1093/llc/fqy085. See also the two-part course "Computer Vision for the Humanities": Daniel van Strien et al., "Computer Vision for the Humanities: An Introduction to Deep Learning for Image Classification (Part 1)," *Programming Historian*, August 17, 2022, accessed March 22, 2024, https://programminghistorian.org/en/lessons/computer-vision-deep-learning-pt1; Daniel van Strien et al., "Computer Vision for the Humanities: An Introduction to Deep Learning for Image Classification (Part 2)," *Programming Historian*, August 17, 2022, accessed March 22, 2024, https://programminghistorian.org/en/lessons/computer-vision-deep-learning-pt2. See furthermore the Special Interest Group "AudioVisual Material in Digital Humanities" (AVinDH) from the Alliance of Digital Humanities Organizations (ADHO), accessed March 22, 2024, https://avindhsig.wordpress.com/; Taylor Arnold et al., "Introduction: Special Issue on AudioVisual Data in DH," *Digital Humanities Quarterly* 15, no. 1 (2021), accessed March 20, 2024, https://www.digitalhumanities.org/dhq/vol/15/1/000541/000541.html.

18 See, for instance, the Sensory Moving Image Archive (SEMIA) project from the University of Amsterdam: https://sensorymovingimagearchive.humanities.uva.nl/, accessed March 22, 2024.

19 Susan Aasman, "Unlocking Multiple Histories of Amateur Media: From Micro- to Macro-Histories," *Screen* 61, no. 1 (2020): 157–166, https://doi.org/10.1093/screen/hjaa011. See also Nicole Braida and Frauke Pirk's chapter in this volume, "Teaching Small-Gauge Formats with Digital Methods."

as media historians Susan Aasman and Tom Slootweg rightfully argue, are "notoriously difficult to reduce to neatly categorized units of analysis, because [they do] not necessarily adhere to formal aesthetic or narrative conventions."[20] At the same time, scholars of amateur film and home movies have been trying to unpack some of the complexities of amateur film production through different scholarly methods and disciplines, ranging from visual anthropology and cultural history to film studies and socio-pragmatics.[21] While digital humanities approaches to amateur films have been emerging and seem promising avenues,[22] computer vision methodologies and the Distant Viewing approach have not yet been explored or tested on this type of audiovisual material.

Our project comprises three interrelated objectives. First of all, we aim to provide insights into the Amateur Film Platform as a collection of historical audiovisual materials. We believe the results will be beneficial for audiovisual archives, as the Distant Viewing Toolkit has the potential to function as a digital enrichment tool, by generating metadata about the films' content and style, for

20 Susan Aasman and Tom Slootweg, "'Pure Information, Not the Real Thing': Digital Hermeneutics and Nelson Sullivan's Videographic Legacy (1983–1989)," this volume.

21 Cf. Richard Chalfen, *Snapshot Versions of Life* (Bowling Green, OH: Bowling Green State University Popular Press, 1987); Patricia R. Zimmermann, *Reel Families: A Social History of Amateur Film* (Bloomington: Indiana University Press, 1995); Roger Odin, ed., *Le film de famille: usage privé, usage public* (Paris: Méridiens Klincksieck, 1995); Susan Aasman, *Ritueel van huiselijk geluk. Een cultuurhistorische verkenning van de familiefilm* (Amsterdam: Het Spinhuis, 2004); Alexandra Schneider, *Die Stars sind wir: Heimkino als filmische Praxis* (Marburg: Schüren, 2004); Martina Roepke, *Privat-Vorstellung: Heimkino in Deutschland vor 1945* (Hildesheim: G. Olms, 2006); Ryan Shand, "Amateur Cinema: History, Theory and Genre (1930–80)" (PhD diss., University of Glasgow, 2007); Karen L. Ishizuka and Patricia R. Zimmermann, eds., *Mining the Home Movie: Excavations in Histories and Memories* (Berkeley: University of California Press, 2008); Laura Rascaroli, Gwenda Young, and Barry Monahan, eds., *Amateur Filmmaking: The Home Movie, the Archive, the Web* (New York: Bloomsbury Publishing, 2014); Charles Tepperman, *Amateur Cinema: The Rise of North American Movie Making, 1923–1960* (Oakland, California: University of California Press, 2015); Annamaria Motrescu-Mayes and Susan Aasman, *Amateur Media and Participatory Cultures: Film, Video, and Digital Media* (London: Routledge, 2019); Masha Salazkina and Enrique Fibla-Gutiérrez, eds., *Global Perspectives on Amateur Film Histories and Cultures* (Bloomington: Indiana University Press, 2021). For amateur video histories, see, among others, James M. Moran, *There's No Place like Home Video* (Minneapolis: University of Minnesota Press, 2002); Diego Cavallotti, "L'audiovisivo analogico della quotidianità. Discorsi, pratiche e testi del cinema e del video amatoriale tra gli anni Settanta e gli anni Novanta in Italia" (PhD diss., University of Udine, 2017); Tom Slootweg, "Resistance, Disruption and Belonging: Electronic Video in Three Amateur Modes" (PhD diss., University of Groningen, 2018).

22 See, for instance, the thematic dossier by Charles Tepperman, "The Complex Materiality of Amateur Cinema Research: Texts, Archives and Digital Methods – Introduction," *Screen* 61, no. 1 (2020): 119–123, https://doi.org/10.1093/screen/hjaa006.

example.[23] Secondly, the project aims to explore the potential of the Distant Viewing approach and toolkit for scholarly purposes, namely for viewing digitized collections of amateur films and videos at scale, specifically for the analysis of formal, stylistic, and aesthetic patterns or changes over time. Thirdly, by doing so, the project more generally aims to contribute to the understanding of the history of amateur film as a cultural practice in the Netherlands. It should be noted that the analysis that empirically informed this chapter was of an *exploratory* nature and based only on a small sample of data (as will be further explained below). Instead of providing a comprehensive analysis of all films and videos published on the Amateur Film Platform, it forms a necessary first step toward a larger endeavor, which requires more resources and infrastructural support.[24]

In the following section, we present our case study and describe the various (iterative) steps and processes involved in "distant viewing" the Amateur Film Platform: from the selection and collection to the analysis, visualization, and interpretation of the data. Utilizing the framework of digital hermeneutics, we critically evaluate how the selected data, tools, and algorithms have impacted the research process and results. In the conclusion, we reflect on the main insights the project brought us and present avenues for future research involving digital approaches to historical amateur film collections and audiovisual materials more broadly.

Distant Viewing the Amateur Film Platform

How could the Distant Viewing approach benefit the study of large amateur film collections? How could the Distant Viewing Toolkit be used as a means to extract, aggregate, and visualize certain semantic metadata and features, such as color use, shot length, object detection, and camera movement in relation to the type of amateur medium or technological carrier used for the recording? How could it be used to analyze any historical, aesthetic, and technological changes in the audiovisual materials from the Amateur Film Platform collection? We were specifically interested in exploring possible patterns in the relationship between different amateur media or technological carriers (e.g., film, video, digital media)

23 Cf. Nanne van Noord et al., "Automatic Annotations and Enrichments for Audiovisual Archives," in *Proceedings of the 13th International Conference on Agents and Artificial Intelligence – Volume 1: ARTIDIGH* (Vienna: SciTePress, 2021), 633–640, https://doi.org/10.5220/0010387706330640.

24 Due to limited resources, access to the Amateur Film Platform collection was constrained. Thanks to the help of Netherlands Institute for Sound & Vision conservator Valentine Kuypers, we were able to select a sample of 100 films for a first test and analysis with the Distant Viewing Toolkit.

on the one hand, and any formal, stylistic, and aesthetic changes on the other, including color versus black-and-white, silent versus sound, shot length, shot type (e.g., total shot, medium shot, close-up), camera movement, use of titles, and setting of recording (indoor versus outdoor).

In conducting our research, we took inspiration from previous projects by the Distant Viewing Lab that implemented the Distant Viewing approach and use of computer vision algorithms for the study of digital collections, such as the Photogrammar and ADDI projects.[25] We also took inspiration from various other digital humanities projects and digital film historical scholarship, including Lev Manovich's pioneering work on cultural analytics, which paved the way for the analysis of large cultural datasets through visualization techniques and data science methods. Cultural analytics allows us, Manovich argues, "to study the patterns, trends, and dynamics" of culture at scale.[26] While Manovich focuses on the analysis of *contemporary* culture, in particular social media, the cultural analytics approach offers a framework for thinking about cultural-historical data as well. Cultural historians and digital humanities scholars Melvin Wevers and Thomas Smits make a plea for a "visual digital turn" and use of computer vision and convolutional neural networks to explore both "the content (what is represented) and the style (how is it represented) of images."[27] Other inspiring projects include the project "Me and Myself: Tracing First Person in Documentary History in AV-Collections" (M&M), led by Susan Aasman, which aimed to find tropes of self-representation in historical Dutch first-person documentary films through digital video annotation,[28] and the Sensory Moving Image Archive (SEMIA) project,

25 For the Photogrammar and ADDI projects, see https://photogrammar.org/maps and https://github.com/distant-viewing/addi, accessed March 22, 2024. Other projects which utilized the Distant Viewing approach include Taylor Arnold, Lauren Tilton, and Justin Wigard, "Automatic Identification and Classification of Portraits in a Corpus of Historical Photographs," in *CEUR Workshop Proceedings* (Computational Humanities Research Conference, Antwerp, Belgium, December 2022), 25–35, accessed March 20, 2024, https://ceur-ws.org/Vol-3290/short_paper5571.pdf; Taylor Arnold, Lauren Tilton, and Annie Berke, "Visual Style in Two Network Era Sitcoms," *Journal of Cultural Analytics* 4, no. 2 (2019), https://doi.org/10.22148/16.043. For a comprehensive list, see the Distant Viewing Lab website: https://distantviewing.org/, accessed March 22, 2024.

26 Manovich, *Cultural Analytics*, 14.

27 Wevers and Smits, "The Visual Digital Turn," 195.

28 Susan Aasman et al., "Tales of a Tool Encounter: Exploring Video Annotation for Doing Media History," *VIEW Journal of European Television History and Culture* 7, no. 14 (2018): 73–87, https://doi.org/10.18146/2213-0969.2018.jethc154. For more information about the CLARIAH pilot project, see "Me and Myself: Tracing First Person in Documentary History in AV-Collections" (2017–2018), https://mediasuite.clariah.nl/learn/example-projects/me-and-myself-tracing-first-person-in-documentary-history-in-av-collections-m-and-m, accessed March 22, 2024.

which explored new ways of searching digitized audiovisual collections based on non-semantic descriptors, such as color, shape, movement, and texture.[29] An interface that similarly enabled the exploration and search of digital film collections based on color was recently implemented in the Timeline of Historical Film Colors, developed by film historian Barbara Flückiger and her team at the University of Zurich, Switzerland. In the related FilmColors project, a digital humanities approach was used to analyze historical film colors on the basis of both technological and aesthetic changes.[30]

We too are interested in the relationship between technological and aesthetic changes from a long-term historical perspective. Utilizing the Distant Viewing approach, one of our initial questions was whether computer vision algorithms could be used to detect the type of amateur medium or technological carrier (film, video, digital media) and even the specific format (e.g., 16mm, 9.5mm, 8mm film) used for recording the amateur film or video. More specifically, we wanted to explore the following aspects:

1. Whether films made on different technological carriers can be distinguished from each other based on specific formal, stylistic, or aesthetic qualities.
2. Whether films made in different time periods can be distinguished from each other based on specific formal, stylistic, or aesthetic qualities.
3. Changes in time (e.g., from black-and-white to color, from silent to sound, shot type and length, use of post-production techniques).
4. Continuities in time (e.g., topics, mise-en-scène, genre, shot type and length, use of post-production techniques).
5. Other aesthetic or narrative tropes (e.g., forms of self-representation, the appearance of amateur film and video cameras, medium-specific characteristics).

29 See https://sensorymovingimagearchive.humanities.uva.nl/, accessed March 22, 2024; Eef Masson, Christian Gosvig Olesen, Nanne van Noord, and Giovanna Fossati, "Exploring Digitised Moving Image Collections: The SEMIA Project, Visual Analysis and the Turn to Abstraction," *Digital Humanities Quarterly* 14, no. 4 (2020), accessed March 20, 2024, http://digitalhumanities.org/dhq/vol/14/4/000497/000497.html.

30 Barbara Flückiger, "A Digital Humanities Approach to Film Colors," *The Moving Image* 17, no. 2 (2017): 71–94, https://doi.org/10.5167/uzh-151113. For the Timeline of Historical Film Colors project, see https://filmcolors.org/, accessed March 22, 2024.

Step 1: Data Selection and Collection

Since the collection presented at the Amateur Film Platform was already digitized,[31] our first step was to collect all the digitized films and videos, so that the collection could be aggregated, analyzed, and visualized by means of the Distant Viewing Toolkit. For various reasons, however, it was not possible to get direct access to all the digital files. The files existed in multiple different digital formats (XML, AVI, MP4, etc.) and were stored in different locations and servers. Moreover, it would require permission from all the archival institutions represented and all the persons whose amateur collections are visible on the platform to re-use their collections for research purposes. We therefore worked with a sample of 100 films and videos, limited to items from the Netherlands Institute for Sound & Vision (NISV), which kindly permitted us to work with this selection of films and metadata from their collection. We felt that limiting the sample data to only the NISV archival collection was justified, as the NISV collection presented the majority of the archival films uploaded to the platform.[32] At the same time, like the other archival institutions represented on the platform, the NISV amateur collection is predominantly film-based. At her presentation at the "e-nedits" meetings in 2020, the NISV conservator of amateur film collections, Valentine Kuypers, mentioned a noticeable peak in film-based materials from the 1930s and late 1950s–1970s in the collection of the Amateur Film Platform (Figure 3).[33] As a result, analogue and digital video formats, like VHS, Betamax, Video2000, Hi8, and MiniDV, which used to be popular amateur recording technologies in the 1980s–1990s, were underrepresented in the collection and on the platform in general. We have tried to circumvent this bias and compensate for the temporal gap by including data from other technological carriers and time periods in the sample. This choice was motivated by our research questions and interest in the study of patterns of historical, technological, and aesthetic changes in amateur media productions.

31 The majority of the films from the NISV collection were digitized as part of the Images for the Future (2007–2014) digitization project, which aimed to digitize, preserve, and make accessible a large part of Dutch audiovisual heritage collections: https://www.beeldengeluid.nl/en/knowledge/projects/images-future, accessed March 22, 2024.

32 At the time of analysis, the NISV collection comprised 2,388 of the 8,491 total items, followed by the Frisian Film Archive (1,129 items) and Groningen Audio Visual Archive (1,093 items); the Community Uploads make for 3,024 items on the platform. A large part of the NISV collection on the Amateur Film Platform stemmed from the Small-Gauge Museum ("Smalfilmmuseum") that was initiated by the late conservator Henk Verheul in 1985 and hosted by the NISV since 2006.

33 Valentine Kuypers, "Amateurfilm Platform" (presentation at the "e-nedits" meetings, INEDITS Amateur Films/Memory of Europe, October 13, 2020), accessed March 22, 2024, http://en.inedits-europe.org/News/Meetings/2020-the-e-nedits-year.

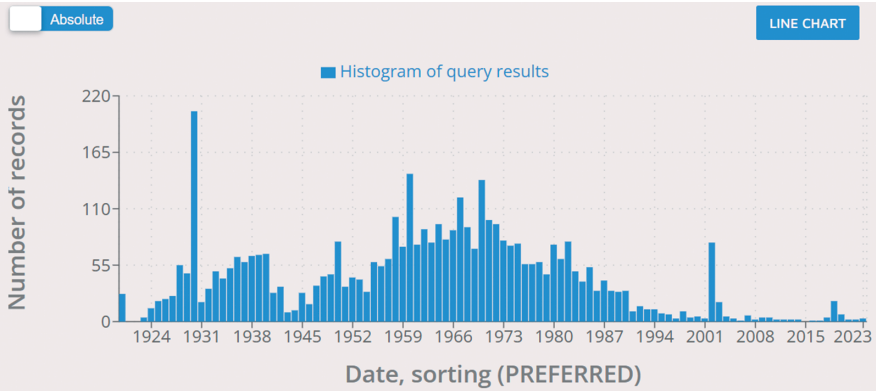

Figure 3: Visualization of the NISV amateur film collection, indicating a bias toward film-based materials from the 1930s and late 1950s–1970s. Source: CLARIAH Media Suite.

In general, we selected our sample of 100 films and videos based on two main criteria and variables: (1) *Time period*: audiovisual items from different decades were selected to enable us to study historical changes; (2) *Medium*: audiovisual items from different amateur media and technological carriers were selected to enable us to study technological changes. For both criteria, we relied on the existing metadata provided by the Amateur Film Platform's interface. As a consequence, items with no specific time period provided in the metadata, and so not appearing in the TIME PERIOD category, were excluded from the selection. The same applied to any item not appearing in the CARRIER category. On the basis of our criteria detailed above, at least two items per decade and per medium were selected. The idea was that this would result in a sample representing all technological carriers (e.g., 35mm, 16mm, 9.5mm, 8mm, Super 8, VHS, MiniDV) and time periods (1900s–2000s). The titles were randomly selected based on their appearance in the Amateur Film Platform search engine (usually alphabetically ordered by film title, sometimes in alphabetical reverse order). In a case where the first two entries were made by the same filmmaker, the following item produced by a different filmmaker was selected to create more diversity in the sample dataset. Where multiple decades were listed in an item's metadata, the item was selected on the basis of the oldest decade: thus a film whose metadata indicated that it included footage of both the 1920s and 1930s was selected as an item representing the 1920s period in our dataset.[34]

34 See, for example, the film item Familiebijeenkomst ter gelegenheid van bezoek uit de vs (Bram Sluis, 1920s-1930s), https://www.amateurfilmplatform.nl/films/sluis-familiefilms-familiebijeenkomst-ter-gelegenheid-van-bezoek-uit-de-vs, accessed October 31, 2023. Archived via: https://web.

The sample dataset was composed in the form of a spreadsheet with various metadata fields to facilitate the aggregation process. The spreadsheet included the following columns: medium (A), time period (B), title of the item (C), file name (D), link to database NISV (E), internal catalog number (F), and URL to Amateur Film Platform (G). The links to the NISV database and internal catalog number were used to facilitate communication with the archive; the URL to the Amateur Film Platform was used for data verification in the analysis phase and for mining relevant metadata, such as content description. After the 100 items had been collected, they were processed by converting the films and videos into a series of still images or frames, so that they could be analyzed by the Distant Viewing Toolkit. In addition to the medium and time period as "source" variables, we selected the following "target" variables: shot length, shot type, setting of recording, color, titles, and camera movement. For measuring these variables and analyzing their relations vis-à-vis the item's medium and time period, three different algorithms were used by the Distant Viewing Toolkit:

1. *Shot boundary detection*:[35] this algorithm helps to detect the number of cuts of an item based on changes in the frame and so helps define the average or median shot length per item (Figure 4).
2. *Image segmentation algorithm*:[36] this algorithm detects objects within the frame, but also backgrounds. Background detection helps to analyze whether a film or scene was recorded indoors or outdoors.
3. *Face detection algorithm*:[37] this algorithm detects people's faces appearing in the frame and, based on this, can be used to define the type of shot (e.g., close-up, medium, total shot).

archive.org/web/20230608061847/https://www.amateurfilmplatform.nl/films/sluis-familiefilms-familiebijeenkomst-ter-gelegenheid-van-bezoek-uit-de-vs.

35 Jakub Lokoč et al., "A Framework for Effective Known-Item Search in Video," in *Proceedings of the 27th ACM International Conference on Multimedia* (New York: Association for Computing Machinery, 2019), 1777–1785, https://doi.org/10.1145/3343031.3351046.

36 Holger Caesar, Jasper Uijlings, and Vittorio Ferrari, "COCO-Stuff: Thing and Stuff Classes in Context," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE Computer Society, 2018), 1209–1218, https://doi.org/10.1109/CVPR.2018.00132.

37 Qiong Cao et al., "VGGFace2: A Dataset for Recognising Faces across Pose and Age," in *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)* (IEEE Computer Society, 2018), 67–74, https://doi.org/10.1109/FG.2018.00020.

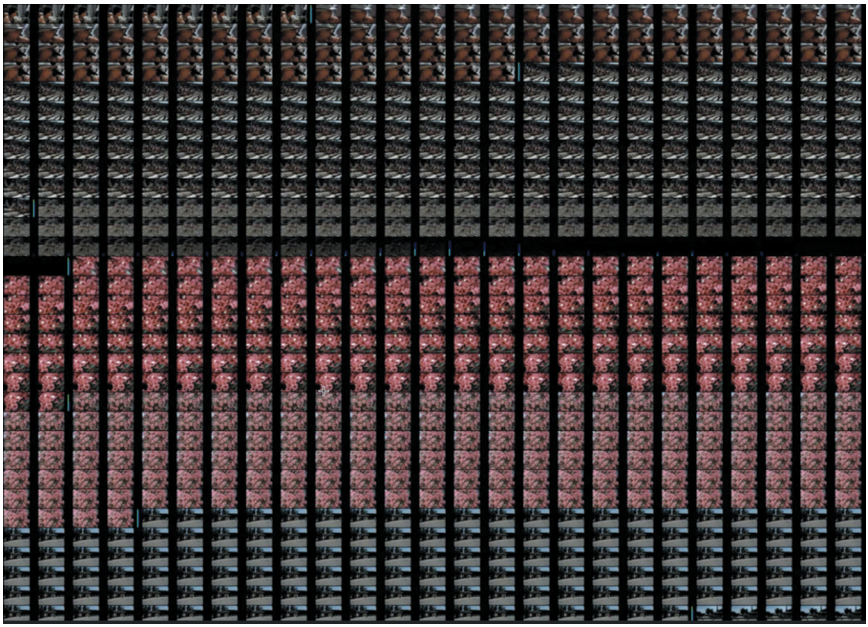

Figure 4: Visualization of the shot boundary detection algorithm being used by the Distant Viewing Toolkit to detect cuts based on frame changes in one of the items from the Amateur Film Platform collection.

Step 2: Data Analysis and Visualization

In various rounds, we analyzed the data sample and tested some of our hypotheses about possible relations between the variables that we had formulated in advance. Specifically, we researched the relationships between medium and time period; shot length in relation to time period and medium; shot type and time period; and setting of recording in relation to time period and medium.[38]

38 In future studies, we would also like to explore the relationships between color and time period; color and medium; film titles and time period; film titles and medium; camera movement and time period; camera movement and medium; topics and time period; and topics and medium.

Medium and Time Period

The first relationship we analyzed was the relation between the technological carrier or medium used and the time period. Based on changes in the emergence and use of certain amateur media technologies in the twentieth century, we hypothesized that film-based items would appear from the 1900s–1980s, electronic video-based items would appear from the 1980s–1990s, and born-digital items from the 1990s–2000s.

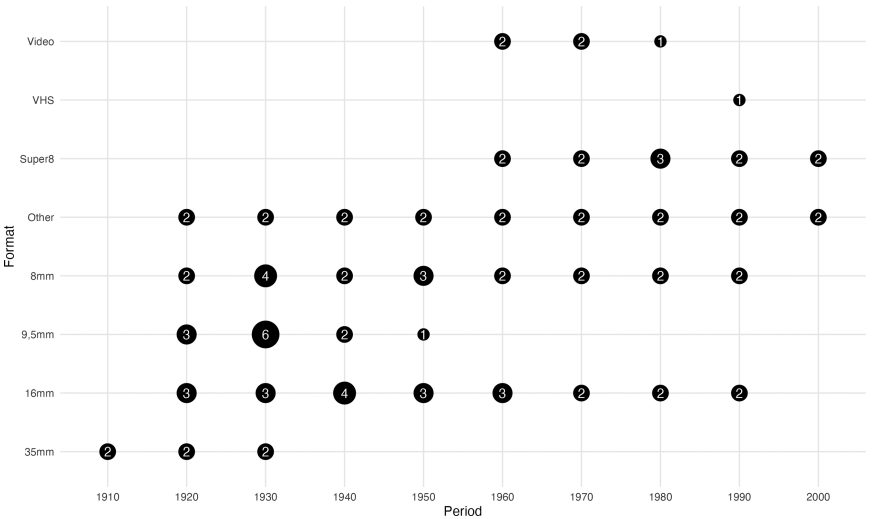

Figure 5: Visualization of the relationship between the medium and time period.

Figure 5 shows how the selected data deviate slightly from the previously mentioned idea of having all technological carriers and time periods equally represented in the sample. Most dots in the graph show the number "2," meaning two audiovisual items were selected for a particular combination of time period and medium, as intended. A variety of reasons can be given in cases where numbers are missing or deviate from this rule. Firstly, a medium has become obsolete or is no longer used by amateurs; for instance, all 35mm films that are included in the Amateur Film Platform collection were made between the 1910s and 1930s. The same was the case for 9.5mm film, which was no longer used as a medium after the 1950s in this sample. Another reason is that a medium was introduced later in time. For instance, Super 8 film was introduced as a technological carrier in 1965, which means no Super 8 films could have been made before the 1960s. The same goes for video-based carriers, like VHS and MiniDV, which were popularized in

the 1980s and 1990s, respectively. Thirdly, the collection itself is limited as it sometimes has only one item from a specific medium available for the selected time period. For instance, there is only one 9.5mm film item from the 1950s, one Video item[39] from the 1980s, and one VHS item from the 1990s. In some cases the dot outnumbers two; for instance, in the 1930s with 9.5mm film (6 items), 8mm film (4 items), and 16mm film (3 items) because, in addition to the selection criteria (time period and medium), a few extra items were selected that addressed specific themes (from exhibitions) or filmmakers that we wanted to include as well (leading to the uneven distribution of technological carriers over time within the sample).

When we look at the graph, we can see that film-based materials are overrepresented in the dataset, also because there are simply more film-based carriers than video-based carriers in the collection. So this was to be expected. More surprising is that Super 8 film items are dominant over video formats in the dataset during the 1980s and 1990s, despite the well-known popularity and widespread use of electronic video in these decades. An explanation for this is the above-mentioned gap in the collection of the Amateur Film Platform, both for electronic video as an amateur medium and for items from the 1980s and 1990s. There are even more films than videos from the 1990s and 2000s, which once again exemplifies the significant bias towards film as a medium and technological carrier in the NISV amateur collection and the collection of the Amateur Film Platform in general.

Shot Length

We used the shot boundary detection algorithm to view and analyze the relations between the shot length and the time period, and between the shot length and the medium.

Shot Length – Time Period

We hypothesized that film-based productions of the first half of the twentieth century would contain more cuts hence a shorter median shot length.[40] This hypoth-

39 See footnote 11.

40 We have chosen the *median* shot length instead of the *average* shot length, following previous research on shot lengths. See Nick Redfern, "The Log-Normal Distribution is Not an Appropriate Parametric Model for Shot Length Distributions of Hollywood Films," *Digital Scholarship in the Humanities* 30, no. 1 (2015): 137–151, https://doi.org/10.1093/llc/fqs066.

esis was informed by the fact that film was expensive, especially in the early twentieth century, so amateur filmmakers would be inclined to use film sparingly and therefore limit the duration of the recording. Apart from costs, the technology of the amateur film camera was a constraining factor: the first spring-based amateur film cameras of the 1920s and 1930s would allow for a maximum constant shooting of approximately 35–40 seconds per shot, whereas electromotor-driven cameras introduced in the 1960s could make longer duration shots.[41]

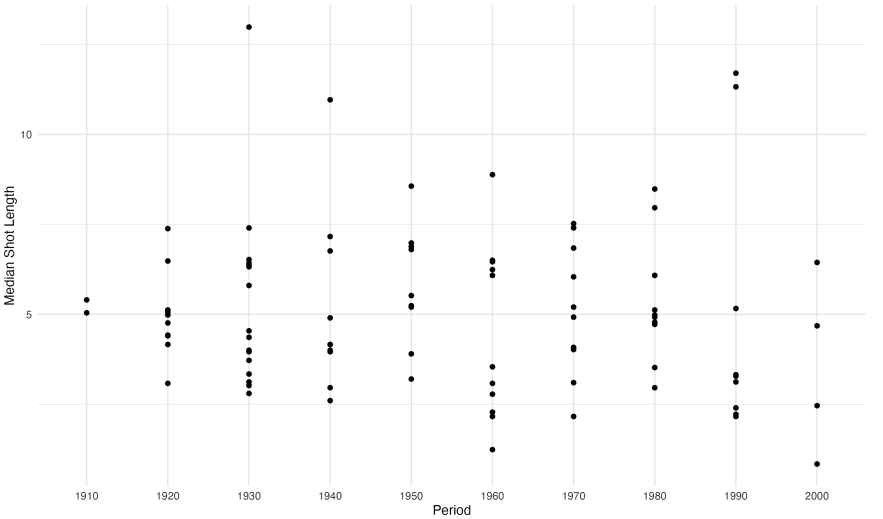

Figure 6: Visualization of the relationship between the time period (in decades) and the median shot length (in seconds).

Figure 6 shows no clear trend or changes in the relationship between the time period and median shot length. Most films and videos from the sample have a median shot length between 2.5 and 7.5 seconds. The figure shows a few outliers. The film with the longest median shot length – 12.98 seconds – was the film AUTOTOCHT NAAR ZWITSERLAND/ITALIË/OOSTENRIJK 1937, made by the amateur filmmaker

41 A spring-driven Ciné-Kodak from the 1930s, for instance, allowed for running 14 to 16 feet of film. See Eastman Kodak Company, *Instructions for Use of the Ciné-Kodak, Model K* (Rochester, New York, ca. 1930), 19. Electromotor-driven cameras no longer had to be manually wound, so could potentially record one take per film roll. Interestingly, manually driven cameras, such as the first Ciné-Kodak camera (model A) from 1923 and Pathé Baby 9.5mm film cameras from 1923–1927, similarly allowed for continuous recording, as their operation was not limited by the spring-drive mechanism.

J. M. Le Grand in 1937.[42] A close analysis of this item reveals that the long median shot length was partly the result of the use of title cards and various jump cuts not detected by the algorithm. Another outlier is the amateur video HUWELIJKSBUS: JAN WIL JE MET ME TROUWEN (Emilia van der Meer, 1993), recorded on VHS, which includes many long takes and so provides a median shot length of 11.32 seconds.[43] This example shows the possible impact of the technological affordance of video cameras to make significantly longer recordings, up to several hours rather than minutes, compared to regular amateur film cameras.[44]

Shot Length – Medium

When looking at the relationship between the shot length and the medium, however, it becomes clear how the technological affordances of the medium do not necessarily determine actual user practices. We hypothesized that because of video's ability to make multiple hours of recordings per tape, there would be fewer cuts in amateur video productions compared to film-based amateur productions. However, if we look at Figure 7, there is no clear trend that indicates this. In fact, there is one noticeable outlier from the sample – the video PROGRESSION (2000) from the amateur film- and video-maker Cor Lievendag[45] – which presents a rather low median shot length (0.49 mu).

Shot Type

We used the face detection algorithm to define the medium height of the face within the frame, which helps to define the shot type: close-up, medium, or total shot. In a close-up shot type, a high percentage of the frame will be filled with the detected face, while in a total shot type, a low percentage of the frame will be filled. Based on these definitions, we were able to view and analyze the relations between the shot type and the time period.

42 See https://www.amateurfilmplatform.nl/films/autotocht-naar-zwitserland-italie-oostenrijk-1937-1, accessed October 31, 2023. Archived via: https://web.archive.org/web/20230402080614/https://www.amateurfilmplatform.nl/films/autotocht-naar-zwitserland-italie-oostenrijk-1937-1.

43 See https://www.amateurfilmplatform.nl/films/huwelijksbus-jan-wil-je-met-me-trouwen-emilia, accessed October 31, 2023. Archived via: https://web.archive.org/web/20230322041545/https://www.amateurfilmplatform.nl/films/huwelijksbus-jan-wil-je-met-me-trouwen-emilia.

44 Moran, *There's No Place like Home Video*, 41; Slootweg, "Resistance, Disruption and Belonging," 213.

45 See https://www.amateurfilmplatform.nl/films/progression, accessed October 31, 2023.

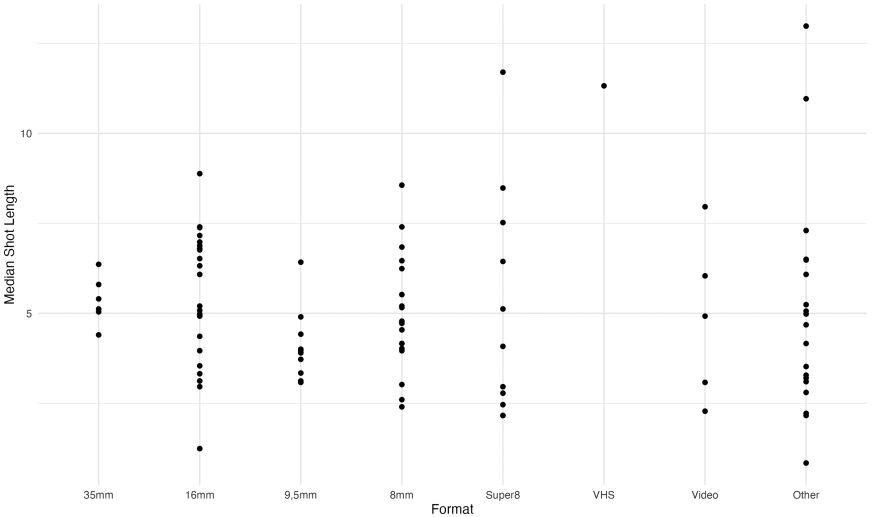

Figure 7: Visualization of the relationship between the medium (per format) and the median shot length (in seconds).

Shot Type – Time Period

Our hypothesis for the relation between the shot type and time period was that close-ups would be increasingly used over time. One of the possible reasons for this aesthetic shift is changes in the affordances of amateur recording technologies. In the 1920s and 1930s, for instance, it was technically more difficult to make close-ups without the risk of recording unsharp images. With early film cameras, one had to manually measure the distance between the lens and the subject in order to set the focus. The later emergence of autofocus functionalities, the possibility of electronic playback for video, and the live preview mode of digital cameras largely eliminated these risks. From the 1960s onwards, we also expected close-ups to become more prominent due to the introduction of the zoom lens, first on film cameras and later also on video recording technologies.

Figure 8 shows a certain trend towards a higher median size of faces based on proportions of the frame height, which could indicate an increase of close-ups as shot type. Items from the 1920s show a small percentage of close-ups with values between around 5 and 12.5, whereas items from the 2000s indicate a larger proportion of the frame being filled with the subject, with values between around 10 and 35. At the same time, however, items with values around 5 continue to be found between the 1920s and 1990s, so films and videos with many medium shots or total shots continue to be made in more recent amateur media productions.

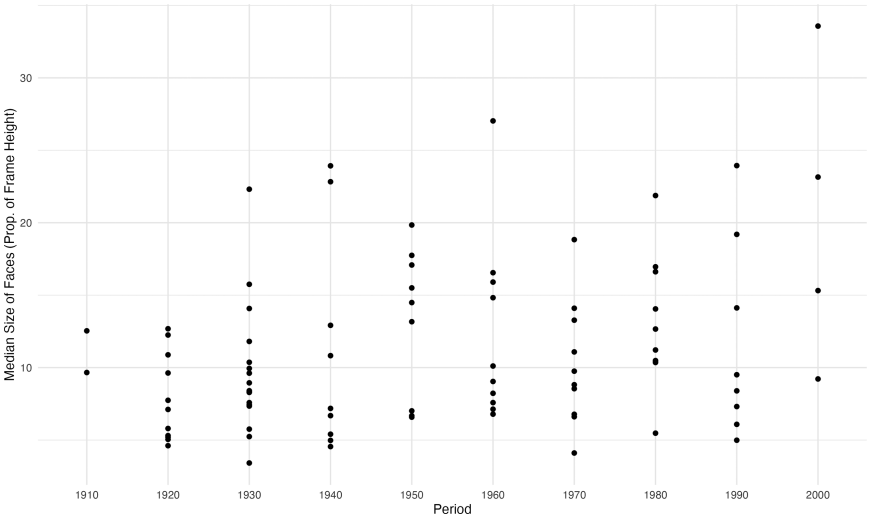

Figure 8: Visualization of the relationship between the time period (in decades) and the type of shot (in median size of faces, properties of the frame height).

Setting of Recording

We used the image segmentation algorithm to define the setting of recording (indoor versus outdoor), which helped us to view and analyze the relations between the setting and the time period, and between the setting and the medium.

Setting – Time Period

The relation between the setting and the time period could signal an aesthetic change from amateur recordings in the first half of the twentieth century to those in the second half. We assumed there would be more outdoor recordings in amateur films from the first half of the twentieth century because of the relatively low light sensitivity of early small-gauge film material. In the 1920s and 1930s, for instance, most reversal films had a speed around 10 ASA, whereas this changed to films with higher speeds ranging from 25–160 ASA and more in the post-war years.[46]

46 Alan Kattelle, *Home Movies: A History of the American Industry, 1897–1979* (Nashua, NH: Transition Pub., 2000), 215, 333–334. See also: https://www.britannica.com/technology/motion-picture-technology/Film#ref508443, accessed August 19, 2024.

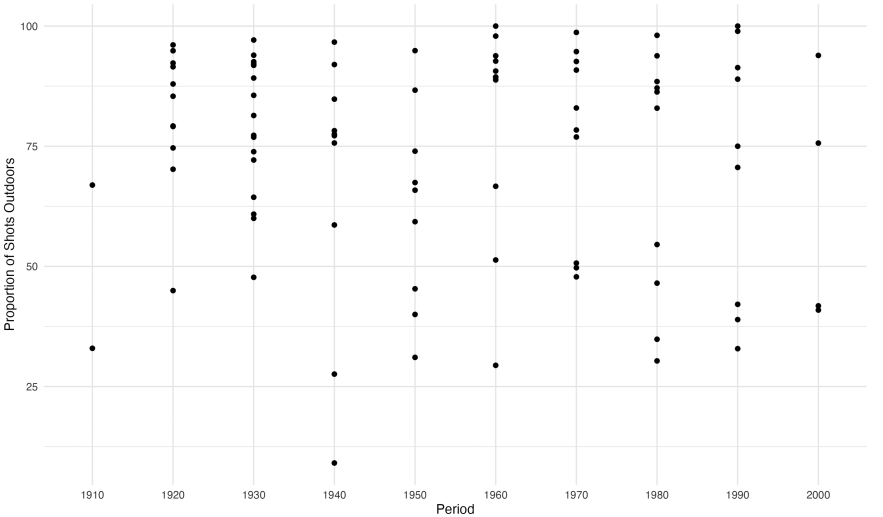

Figure 9: Visualization of the relationship between the time period (in decades) and the setting of recording (based on the proportion of outdoor shots detected by the algorithm).

Figure 9 shows no clear trend, however. The majority of films from our sample seem to have been recorded outdoors, with a few exceptions in each time period, most notably one item from the 1940s, KLEUTERKLAS, HUISHOUDSCHOOL (H. de Jong-Kleinweg de Zwaan, 1940-1941), where the subject consisted entirely of the preparation and serving of a formal meal in an indoor dining room.[47] In general, more variations in recording settings seem to have occurred from the 1940s onwards, possibly due to the above-mentioned changes in the technological affordances of the recording medium.

Setting – Medium

In relation to the medium used as a variable, we assumed there would be more indoor recordings in video-based amateur productions than in film-based amateur productions.[48] The first electronic video cameras that appeared in the 1960s and 1970s

47 See https://www.amateurfilmplatform.nl/films/kleuterklas-huishoudschool-1940-1941-2, accessed October 31, 2023. Archived via: https://web.archive.org/web/20231003132327/https://www.amateurfilmplatform.nl/films/kleuterklas-huishoudschool-1940-1941-2.
48 Moran, *There's No Place like Home Video*, 41.

were directly connected to a stationary video recorder.[49] Only later, portable video recording technologies were introduced that allowed for more flexible forms of recording, both indoors and outdoors.

Figure 10 does not indicate a significant difference between the medium and the setting of recording or between film- and video-based amateur productions specifically. In general, however, it seems that the majority of 8mm, 9.5mm and Video items were recorded outdoors (50% or higher proportions of shots outside), whereas for amateur films recorded on 35mm, 16mm, and Super 8 more variations occurred, hence these technological carriers included items with relatively more indoor shots (below 50%).

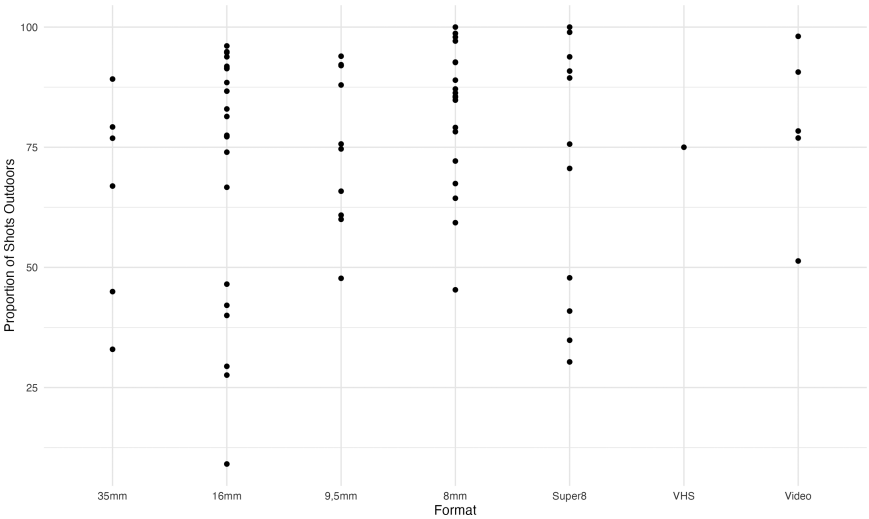

Figure 10: Visualization of the relationship between the medium (per format) and the setting of recording (based on the proportion of shots outdoors).

Step 3: Data Interpretation and Criticism

How to interpret the analyzed relationships from the previous step? What new insights do they provide? How do they reflect patterns or trends in amateur filmmaking in the Netherlands? And how can we critically reflect on these results and the ways in which they were produced by the selected algorithms? In this

49 Van der Heijden, "Hybrid Histories," 178–179.

final step, we draw on the framework of *digital hermeneutics*, defined as a "concept that enables historians to critically reflect on the various interventions of digital research infrastructures, tools, databases, and dissemination platforms in the process of thinking, doing and narrating history."[50] The digital hermeneutics framework helps us to interpret the data and reflect on how the Distant Viewing approach and toolkit have shaped our research practices and findings.

Data Bias, Patterns, and Source Criticism

For the interpretation of the results, it is important to reflect first of all on the notion of *data representation*: how representative are the results and what do they tell us? Since the collection of the Amateur Film Platform indicated a strong bias towards film-based materials, we tried to balance this out by selecting items from multiple time periods and technological carriers. Constructing a sample dataset based on these criteria helped us to study certain patterns and changes in the collection, as we have seen in the previous step. Yet, it also led to a new form of bias, which prevented the sample from being representative, either for the collection of the Amateur Film Platform or for Dutch amateur media productions in general. Based on the small sample of data we analyzed, it is therefore difficult to make valid statements about relationships, trends, and patterns in amateur media collections and the possible historical, technological, or aesthetic changes they reflect. The value of the present study lies, we would argue, on the methodological level and the various questions it raises about using digital methods for the study of historical amateur media collections.

Such questions pertain to the digital transformation of archival materials. While digital methods may create new biases, they can also help to make existing biases and their underlying "politics of digitization" in archival databases more explicit.[51] In relation to the Amateur Film Platform collection, the bias towards film-based collections largely stems from collection policies, which generally prioritized film over video as cultural heritage objects in digitization projects.[52] This also

50 Andreas Fickers, Juliane Tatarinov, and Tim van der Heijden, "Digital History and Hermeneutics – Between Theory and Practice: An Introduction," in *Digital History and Hermeneutics*, ed. Andreas Fickers and Juliane Tatarinov (Berlin: De Gruyter, 2022), 6–7, https://doi.org/10.1515/9783110723991-001.

51 Gerben Zaagsma, "Digital History and the Politics of Digitization," *Digital Scholarship in the Humanities* 38, no. 2 (2023): 830–851, https://doi.org/10.1093/llc/fqac050.

52 Frank Holthuizen, Joseph Wachelder, and Tim van der Heijden, "Amateurfilm in Limburg, Limburg in Amateurfilm," in *Publications de la Société Historique et Archéologique dans le Limburg*, vol. 156 (Maastricht: Koninklijk Limburgs Geschied- en Oudheidkundig Genootschap LGOG, 2021), 229–281.

touches upon the importance of *digital source criticism* for reflecting on provenance and digitization processes: when and how have the source materials been digitized, and how has this influenced the results? The majority of the digitized films that we used in our sample were digitized in standard definition (SD) resolution, while more recently digitized items in the collection have been scanned in high definition (HD) or even 2K image resolution.[53] The question of how differences in digitization practices and scan resolutions affect the application of the Distant Viewing Toolkit algorithms would be an interesting avenue for further exploration and comparative testing, but was beyond the scope of the present study.

In addition to questions of archival politics and practices of digitization, another point of reflection is the use of *metadata* and the way they have been structured, and hence shaped our research. As Johanna Drucker reminds us, metadata as a "system of naming, organizing, and classifying" of materials is not neutral but influences how archival items are searched and researched to a great extent.[54] In the case of the Amateur Film Platform, in fact, we have been heavily relying on the platform's metadata for selecting the items and constructing our data sample. Instead of making categorizations ourselves based on certain search criteria, we took the existing categories provided by the platform to select items from the different time periods and technological carriers. Although this enabled us to quickly select various potentially relevant items, it also entailed two large disadvantages as discussed: mislabeling and the exclusion of relevant items.

Data Conversion, Algorithms, and Tool Criticism

In addition to archival politics, digitization practices, and metadata models, our results were shaped by the affordances and constraints of the Distant Viewing approach. Applying questions of *algorithmic criticism* and *tool criticism*,[55] we can reflect on how the toolkit and its algorithms have shaped or structured our analysis and interpretation of the data. One way the toolkit restructures the data is by

53 Cf. Franziska Heller, *Update!* (Paderborn: Wilhelm Fink Verlag, 2020), https://doi.org/10.30965/9783846764602. See also Franziska Heller and Ulrich Ruedel, "Pursuing Film History with Digital Images: Towards Visual Literacy in the Age of AI and Social Media," this volume.

54 Johanna Drucker, *The Digital Humanities Coursebook: An Introduction to Digital Methods for Research and Scholarship*, 1st ed. (Abingdon, Oxon; New York: Routledge, 2021), 53.

55 Stephen Ramsay, *Reading Machines: Toward an Algorithmic Criticism* (Urbana, IL: University of Illinois Press, 2017), https://doi.org/10.5406/illinois/9780252036415.001.0001; Marijn Koolen, Jasmijn Van Gorp, and Jacco van Ossenbruggen, "Toward a Model for Digital Tool Criticism: Reflection as Integrative Practice," *Digital Scholarship in the Humanities* 34, no. 2 (2019): 368–385, https://doi.org/10.1093/llc/fqy048.

converting the audiovisual items into a series of still images or frames, so that they can be processed by the computer vision algorithms. While this step is necessary to enable us to computationally analyze audiovisual materials at scale, it also implies the transformation of the data and removal of some of the original qualities of the film and video items, including their temporal dimensions and, where applicable, their sound layer.[56] Although there are no suitable algorithms at the moment of writing that could be used by the Distant Viewing Toolkit to meaningfully analyze the sound layer in addition to the visual layer, this multimodality would be a crucial aspect to take into account in future explorations.

Other current limitations we encountered when applying the Distant Viewing Toolkit's algorithms pertain to certain aesthetic characteristics or medium-specific features. Sometimes, for instance, the toolkit's algorithms had difficulties with the interpretation of film titles. The image segmentation algorithm mistakenly detected some of those with the label "sky." Furthermore, in cases where films made use of fades (i.e., fade-in or fade-out), the shot boundary detection algorithm sometimes encountered difficulties when counting the number of cuts. Such issues can be explained by the fact that the computer vision algorithms that were used have not been trained (yet) on historical audiovisual materials, let alone amateur media productions, so we anticipated they would not work "perfectly" on digitized analogue films and videos. We were actually surprised how well the algorithms worked on even the oldest black-and-white films from the collection from the 1910s and 1920s. In fact, they worked even better on the older material compared to the electronic video items from the 1980s–1990s due to the visual noise and grainy quality of the latter. Correcting these mistakes from the computer vision algorithms prevented our work with the Distant Viewing Toolkit being a straightforward or linear process; it was rather a cyclical or iterative process, in which algorithms usually had to be "adjusted" to achieve more reliable results.

Conclusion

In this chapter, we used the Distant Viewing approach and toolkit to navigate and analyze the Amateur Film Platform collection at scale. In our study, which serves as a pilot project for a larger and more comprehensive study on the application of computer vision algorithms to amateur films and videos as audiovisual materials from archival collections, we wanted to explore possible occurrences of formal, stylistic, or aesthetic patterns in the collection based on changes in periods

56 See also Aasman and Slootweg, "Pure Information, Not the Real Thing," this volume.

of time and technological carriers. In response to our five sub-questions formulated in the beginning of this chapter, we can make the following preliminary conclusions based on the results from our sample analysis:

1. Films made on different technological carriers can, to some extent, be distinguished from each other based on specific formal, stylistic, or aesthetic qualities. There is no correlation between the amateur medium or technological carrier and the shot length. However, there seems to be some kind of relationship between a selection of technological carriers and the setting of recording: 8mm, 9.5mm and Video items were generally recorded outdoors while amateur films recorded on 35mm, 16mm, and Super 8 included indoor recordings as well.
2. Films made in different time periods can, to some extent, be distinguished from each other based on specific formal, stylistic, or aesthetic qualities. While we could not find a clear trend in the relationship between the shot length and time period, there seems to be a trend toward an increased use of medium and close-up shot types over time (higher median size of faces based on proportions of the frame height). No trend could be seen for the relation between time period and the setting of recording, although more variations in indoor and outdoor recordings seem to occur from the 1940s onwards, possibly due to the emergence of more light-sensitive amateur film material and recording equipment that enabled indoor recording practices without the use of external lamps.
3. The most noticeable changes over time were shown in the sample when it comes to changes in shot type, namely an increase of close-ups. Changes in amateur media productions from black-and-white to color, as well as from silent to sound, we also expect to be significant. Possible changes in post-production techniques, like use of title animations and montage, are probably more difficult to detect computationally.
4. The most significant continuities were found in the recording setting: most amateur films and videos from the sample indicate a high proportion of outdoor shots. In addition, the relationship between shot type and time period shows the continuation of films and videos with many medium shots and total shots.
5. No other aesthetic or narrative tropes were found in the present study. However, forms of self-representation and the appearance of amateur recording technologies like film and video cameras within the frame, to be detected via object detection algorithms, could be interesting tropes to explore in follow-up studies.

Distant and Close Viewing: Towards a Hybrid Heuristics

Clearly, Distant Viewing has considerable potential for the study of large datasets of historical amateur media collections. As demonstrated in previous projects of the Distant Viewing Lab, the data-driven approach to audiovisual data does not replace traditional film analytical methods, but rather stimulates or provides the starting point for further research and investigation. The Distant Viewing Toolkit provides a way to zoom in and out on a corpus of audiovisual data and to see possible patterns, deviations, and relations between items, in our case the sample of films and videos from the Amateur Film Platform collection. In particular, the visualization of outliers in the data can be useful as a heuristic instrument. Outliers in the data help to quickly detect those audiovisual items that deviate from a certain trend or norm. In our sample, for example, the high-speed montage video PROGRESSION (2000) stood out because of its low median shot length. Outliers can also be used effectively to find errors or biases in the algorithms and underlying code systems, when a film was mislabeled in a certain category, for instance, or when the algorithm provided an erroneous qualification. The image segmentation algorithm that we used for detecting the setting of recording (indoor versus outdoor), for example, mistakenly analyzed some of the overexposed images with the tag "snow" and hence mislabeled the scene as "outdoors." Such discrepancies illustrate once more that the results from using digital methods should never be taken for granted, but carefully evaluated and interpreted.

The question of how to "view" and "interpret" the data is a challenging one in general. This is especially the case when it comes to amateur-produced materials, which tend to become meaningful not so much through their aesthetic or narrative features but rather in relation to their contexts of production and reception.[57] For the FilmColors research project, Barbara Flückiger argued that one of the pitfalls of quantitative analysis "is its potential to disregard the meaningful context of data occurrences across the body of works studied."[58] While a data-driven approach may indeed enable new possibilities and perspectives on large collections,[59] the use of digital methods should not prevent a complementary, contextualist approach and close analysis for comprehensively understanding historical materials and their specific-

57 Odin, *Le film de famille*, 36; Aasman, *Ritueel van huiselijk geluk*, 73.

58 Flückiger, "A Digital Humanities Approach to Film Colors," 72.

59 In the FilmColors project these new possibilities included the identification of "diachronic aesthetic patterns" of film color aesthetics. See Flückiger, "A Digital Humanities Approach to Film Colors," 72.

ities. The combination of digital methods with close analysis is crucial.[60] In conclusion, we would therefore like to argue for a *hybrid heuristics*, which combines "distant" and "close" viewing, computational and hermeneutic modes, data-driven and case-study-driven forms of analysis, as well as an approach that combines the affordances of digital technologies and methods, such as computer vision, with the critical reflection on their hermeneutic implications, such as biases and data structures.

Next Steps

Potential next steps in applying the Distant Viewing approach for the study of amateur media histories includes the development of more complex analyses, involving more than two variables and algorithms. This would be useful for studying certain aspects of amateur filmmaking as a cultural practice, for instance whether a film was recorded on a tripod or by means of a hand-held camera. Such information could possibly be retrieved computationally by combining the image segmentation and face detection algorithms. Similarly, more complex queries are required for analyzing practices and forms of self-representation (e.g., detecting changes in the ways amateur filmmakers have portrayed themselves in their films and how this has been shaped by the medium used),[61] the appearance of technological objects (e.g., film/video camera), and medium specificity (e.g., recognition of technological carrier on the basis of certain aesthetic and material characteristics, such as film grain, perforation type, manufacturing codes, and other relevant information about the audiovisual material). The detection of the average number of "out of focus" shots per film or video might also be interesting to explore historically.

Another potentially interesting avenue for further exploration is to investigate how the combination of multiple variables and algorithms could shed new light on differences in types of amateur users, such as "family filmmakers" who make films to capture family memories versus "serious hobbyists" who make films as a (technical) hobby.[62] Could we confirm, for example, any patterns indicating that serious hobbyists use more film titles, have a greater variation in shot length and shot type in their films, and prefer 16mm over 8mm film as recording medium? The question of whether Distant Viewing could be used to detect the type of technological carrier

60 See also Casper Tybjerg, Jonatan Bruun Borring, and Luan Nhu Vu, "The Digitization of Silent Films and the Teaching of Film Historiography: Entanglements and Opportunities," this volume.
61 Aasman et al., "Tales of a Tool Encounter." See also Aasman and Slootweg, "Pure Information, Not the Real Thing," this volume.
62 Aasman, *Ritueel van huiselijk geluk*, 39–64; Van der Heijden, "Hybrid Histories," 89–90.

or medium used, such as whether the amateur film was recorded on 16mm, 9.5mm, 8mm, or Super 8, was one of the questions that inspired our project in the first place.[63] At this stage, however, it seems unlikely it would be possible to detect this computationally when no information on or about the perforation is included in the scan, even when doing extensive development and additional training of the algorithms. This is because there are simply too many factors that may influence the scan results and hence the quality of the digitized film, from varieties in the type of film material and circumstances of film production to how and when the film was digitized and by whom.

Nevertheless, we are convinced that there is great potential in the Distant Viewing approach and the use of computer vision methods for the study of historical amateur film collections at scale. Not only for scholarly purposes, but also for audiovisual archives, it may provide new ways to access, navigate, extract, and enrich metadata from digitized collections.[64] For understanding historical amateur film collections and the history of amateur filmmaking as a cultural practice, our project provided a first step towards a larger project with more infrastructural support. Rather than working with a data sample, we would like to fully analyze and internationally compare multiple large collections of amateur films and videos. Such a comprehensive and comparative approach would truly make use of the potential of Distant Viewing for analyzing amateur media collections at scale, thereby making visible various (hidden) patterns, relations, changes, and continuities in the history of amateur filmmaking as a cultural practice from a long-term historical perspective.

63 This question was raised after visiting Jasper Rigole's exhibition "Homeless Movies" at the House of Alijn, Museum of Everyday Life in Ghent, Belgium, in 2016–2017. Rigole is a Belgian visual artist, researcher, and collector of home movies and found footage materials. The exhibition presented Rigole's IICADOM collection, on the basis of which he constructed "an elementary taxonomy of collected memory" that allowed visitors to search and filter the collection of home movies by theme, subject, genre, action, location, object, and formal characteristics. See Jasper Rigole, *Addenda* (Ghent: AraMER, 2015), accessed August 20, 2024, https://www.jubilee-art.org/projects/addenda. For the IICADOM website, see https://iicadom.org/, accessed March 22, 2024.

64 It would also be worthwhile to explore the question of how Distant Viewing could be used for enriching metadata of archival film collections further in relation to film identification processes. Cf. Harold Brown and Camille Blot-Wellens, *Physical Characteristics of Early Films as Aids to Identification*, 2nd ed. (Brussels: FIAF, 2020).

References

Aasman, Susan. “Unlocking Multiple Histories of Amateur Media: From Micro- to Macro-Histories.” *Screen* 61, no. 1 (2020): 157–166. https://doi.org/10.1093/screen/hjaa011.

Aasman, Susan. “Finding Traces in YouTube’s Living Archive: Exploring Informal Archival Practices.” *TMG Journal for Media History* 22, no. 1 (2019): 35–55. https://doi.org/10.18146/tmg.435.

Aasman, Susan. *Ritueel van huiselijk geluk. Een cultuurhistorische verkenning van de familiefilm*. Amsterdam: Het Spinhuis, 2004.

Aasman, Susan, Liliana Melgar Estrada, Tom Slootweg, and Rob Wegter. “Tales of a Tool Encounter: Exploring Video Annotation for Doing Media History.” *VIEW Journal of European Television History and Culture* 7, no. 14 (2018): 73–87. https://doi.org/10.18146/2213-0969.2018.jethc154.

Acland, Charles R., and Eric Hoyt, eds. *The Arclight Guidebook to Media History and the Digital Humanities*. Falmer: REFRAME Books, 2016. Accessed March 22, 2024. https://projectarclight.org/book/.

Arnold, Taylor, Jasmijn Van Gorp, Stefania Scagliola, and Lauren Tilton. “Introduction: Special Issue on AudioVisual Data in DH.” *Digital Humanities Quarterly* 15, no. 1 (2021). Accessed March 20, 2024. https://www.digitalhumanities.org/dhq/vol/15/1/000541/000541.html.

Arnold, Taylor, and Lauren Tilton. *Distant Viewing: Computational Exploration of Digital Images*. Cambridge, MA: MIT Press, 2023.

Arnold, Taylor, and Lauren Tilton. “Distant Viewing Toolkit: A Python Package for the Analysis of Visual Culture.” *Journal of Open Source Software* 5, no. 45 (2020): 1–6. https://doi.org/10.21105/joss.01800.

Arnold, Taylor, and Lauren Tilton. “Distant Viewing: Analyzing Large Visual Corpora.” *Digital Scholarship in the Humanities* 34, issue supplement 1 (2019): i3–i16. https://doi.org/10.1093/llc/fqz013.

Arnold, Taylor, Lauren Tilton, and Annie Berke. “Visual Style in Two Network Era Sitcoms.” *Journal of Cultural Analytics* 4, no. 2 (2019). https://doi.org/10.22148/16.043.

Arnold, Taylor, Lauren Tilton, and Justin Wigard. “Automatic Identification and Classification of Portraits in a Corpus of Historical Photographs.” In *CEUR Workshop Proceedings*, 25–35. Computational Humanities Research Conference, Antwerp, Belgium, December 2022. Accessed March 20, 2024. https://ceur-ws.org/Vol-3290/short_paper5571.pdf.

Brown, Harold, and Camille Blot-Wellens. *Physical Characteristics of Early Films as Aids to Identification*, 2nd ed. Brussels: FIAF, 2020.

Caesar, Holger, Jasper Uijlings, and Vittorio Ferrari. “COCO-Stuff: Thing and Stuff Classes in Context.” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 1209–1218. IEEE Computer Society, 2018. https://doi.org/10.1109/CVPR.2018.00132.

Cao, Qiong, Li Shen, Weidi Xie, Omkar M. Parkhi, and Andrew Zisserman. “VGGFace2: A Dataset for Recognising Faces across Pose and Age.” In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 67–74. IEEE Computer Society, 2018. https://doi.org/10.1109/FG.2018.00020.

Cavallotti, Diego. “L’audiovisivo analogico della quotidianità. Discorsi, pratiche e testi del cinema e del video amatoriale tra gli anni Settanta e gli anni Novanta in Italia.” PhD diss., University of Udine, 2017.

Chalfen, Richard. *Snapshot Versions of Life*. Bowling Green, OH: Bowling Green State University Popular Press, 1987.

Drucker, Johanna. *The Digital Humanities Coursebook: An Introduction to Digital Methods for Research and Scholarship*, 1st ed. Abingdon, Oxon; New York: Routledge, 2021.

Eastman Kodak Company. *Instructions for Use of the Ciné-Kodak, Model K*. Rochester, New York, ca. 1930.

Enticknap, Leo. *Moving Image Technology: From Zoetrope to Digital*. London, New York: Wallflower, 2005.

Fickers, Andreas, Juliane Tatarinov, and Tim van der Heijden. "Digital History and Hermeneutics – Between Theory and Practice: An Introduction." In *Digital History and Hermeneutics*, edited by Andreas Fickers and Juliane Tatarinov, 1–20. Berlin: De Gruyter, 2022. https://doi.org/10.1515/9783110723991-001.

Flückiger, Barbara. "A Digital Humanities Approach to Film Colors." *The Moving Image* 17, no. 2 (2017): 71–94. https://doi.org/10.5167/uzh-151113.

Heller, Franziska. *Update!* Paderborn: Wilhelm Fink Verlag, 2020. https://doi.org/10.30965/9783846764602.

Holthuizen, Frank, Joseph Wachelder, and Tim van der Heijden. "Amateurfilm in Limburg, Limburg in Amateurfilm." In *Publications de la Sociéte Historique et Archéologique dans le Limburg*, vol. 156, 229–281. Maastricht: Koninklijk Limburgs Geschied- en Oudheidkundig Genootschap LGOG, 2021.

Ishizuka, Karen L., and Patricia R. Zimmermann, eds. *Mining the Home Movie: Excavations in Histories and Memories*. Berkeley: University of California Press, 2008.

Kattelle, Alan. *Home Movies: A History of the American Industry, 1897–1979*. Nashua, NH: Transition Pub., 2000.

Koolen, Marijn, Jasmijn Van Gorp, and Jacco van Ossenbruggen. "Toward a Model for Digital Tool Criticism: Reflection as Integrative Practice." *Digital Scholarship in the Humanities* 34, no. 2 (2019): 368–385. https://doi.org/10.1093/llc/fqy048.

Kuypers, Valentine. "Amateurfilm Platform." Presentation at the "e-nedits" meetings, INEDITS Amateur Films/Memory of Europe, October 13, 2020. Accessed March 22, 2024. http://en.inedits-europe.org/News/Meetings/2020-the-e-nedits-year.

Lokoč, Jakub, Gregor Kovalčik, Tomáš Souček, Jaroslav Moravec, and Přemysl Čech. "A Framework for Effective Known-Item Search in Video." In *Proceedings of the 27th ACM International Conference on Multimedia*, 1777–1785. New York: Association for Computing Machinery, 2019. https://doi.org/10.1145/3343031.3351046.

Lupu, Ruxandra. "The Home Movie 4.0: (Co)Creative Strategies for a Tacit, Embodied and Affective Reading of the Sicilian Home Movie Archive." PhD diss., University of Leeds, 2020. Accessed August 30, 2023. https://etheses.whiterose.ac.uk/27966/.

Manovich, Lev. *Cultural Analytics*. Cambridge, MA: MIT Press, 2020.

Manovich, Lev. *The Language of New Media*. Cambridge, MA: MIT Press, 2001.

Masson, Eef, Christian Gosvig Olesen, Nanne van Noord, and Giovanna Fossati. "Exploring Digitised Moving Image Collections: The SEMIA Project, Visual Analysis and the Turn to Abstraction." *Digital Humanities Quarterly* 14, no. 4 (2020). Accessed March 30, 2024. http://digitalhumanities.org/dhq/vol/14/4/000497/000497.html.

Moran, James M. *There's No Place like Home Video*. Minneapolis: University of Minnesota Press, 2002.

Motrescu-Mayes, Annamaria, and Susan Aasman. *Amateur Media and Participatory Cultures: Film, Video, and Digital Media*. London: Routledge, 2019.

Noord, Nanne van, Christian Gosvig Olesen, Roeland Ordelman, and Julia Noordegraaf. "Automatic Annotations and Enrichments for Audiovisual Archives." In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence – Volume 1: ARTIDIGH*, 633–640. Vienna: SciTePress, 2021. https://doi.org/10.5220/0010387706330640.

Odin, Roger, ed. *Le film de famille: usage privé, usage public*. Paris: Méridiens Klincksieck, 1995.

Ramsay, Stephen. *Reading Machines: Toward an Algorithmic Criticism*. Urbana, IL: University of Illinois Press, 2017. https://doi.org/10.5406/illinois/9780252036415.001.0001.

Rascaroli, Laura, Gwenda Young, and Barry Monahan, eds. *Amateur Filmmaking: The Home Movie, the Archive, the Web*. New York: Bloomsbury Publishing, 2014.

Redfern, Nick. "The Log-Normal Distribution is Not an Appropriate Parametric Model for Shot Length Distributions of Hollywood Films." *Digital Scholarship in the Humanities* 30, no. 1 (2015): 137–151. https://doi.org/10.1093/llc/fqs066.

Rigole, Jasper. *Addenda*. Ghent: AraMER, 2015. Accessed August 20, 2024. https://www.jubilee-art.org/projects/addenda.

Roepke, Martina. *Privat-Vorstellung: Heimkino in Deutschland vor 1945*. Hildesheim: G. Olms, 2006.

Salazkina, Masha, and Enrique Fibla-Gutiérrez, eds. *Global Perspectives on Amateur Film Histories and Cultures*. Bloomington: Indiana University Press, 2021.

Schneider, Alexandra. *Die Stars sind wir: Heimkino als filmische Praxis*. Marburg: Schüren, 2004.

Shand, Ryan. "Amateur Cinema: History, Theory and Genre (1930–80)." PhD diss., University of Glasgow, 2007.

Simoni, Paolo. "The Amateur City: Digital Platforms and Tools for Research and Dissemination of Films Representing the Italian Urban Landscape." *The Moving Image: The Journal of the Association of Moving Image Archivists* 17, no. 2 (2017): 111–118. https://doi.org/10.5749/movingimage.17.2.0111.

Slootweg, Tom. "Resistance, Disruption and Belonging: Electronic Video in Three Amateur Modes." PhD diss., University of Groningen, 2018.

Strien, Daniel van, Kaspar Beelen, Melvin Wevers, Thomas Smits, and Katherine McDonough. "Computer Vision for the Humanities: An Introduction to Deep Learning for Image Classification (Part 1)." *Programming Historian*, August 17, 2022. Accessed March 22, 2024. https://programminghistorian.org/en/lessons/computer-vision-deep-learning-pt1.

Strien, Daniel van, Kaspar Beelen, Melvin Wevers, Thomas Smits, and Katherine McDonough. "Computer Vision for the Humanities: An Introduction to Deep Learning for Image Classification (Part 2)." *Programming Historian*, August 17, 2022. Accessed March 22, 2024. https://programminghistorian.org/en/lessons/computer-vision-deep-learning-pt2.

Sui, Patrick. "Review: Distant Viewing Toolkit." *Reviews in Digital Humanities* 2, no. 4 (2021). https://doi.org/10.21428/3e88f64f.01fedc58.

Tepperman, Charles. "The Amateur Movie Database." *Screen* 61, no. 1 (2020): 124–128. https://doi.org/10.1093/screen/hjaa007.

Tepperman, Charles. "The Complex Materiality of Amateur Cinema Research: Texts, Archives and Digital Methods – Introduction." *Screen* 61, no. 1 (2020): 119–123. https://doi.org/10.1093/screen/hjaa006.

Tepperman, Charles. "The Amateur Movie Database: Archives, Publics, Digital Platforms." *The Moving Image: The Journal of the Association of Moving Image Archivists* 17, no. 2 (2017): 106–110. https://doi.org/10.5749/movingimage.17.2.0106.

Tepperman, Charles. *Amateur Cinema: The Rise of North American Movie Making, 1923–1960*. Oakland, California: University of California Press, 2015.

Van der Heijden, Tim. "Technologies of Memory: Amateur Storage Media and Home Movie Practices in the Longue Durée." *Le Temps Des Médias* 39, no. 2 (2022): 141–159. https://doi.org/10.3917/tdm.039.0141.

Van der Heijden, Tim. "Hybrid Histories: Technologies of Memory and the Cultural Dynamics of Home Movies, 1895–2005." PhD diss., Maastricht University, 2018.

Van der Heijden, Tim, and Valentine Kuypers. "Life in Motion: A History of Amateur Film." *Europeana*, April 12, 2023. Accessed August 8, 2024. https://www.europeana.eu/en/exhibitions/life-in-motion/.

Wevers, Melvin, and Thomas Smits. "The Visual Digital Turn: Using Neural Networks to Study Historical Images." *Digital Scholarship in the Humanities* 35, no. 1 (2020): 194–207. https://doi.org/10.1093/llc/fqy085.

Zaagsma, Gerben. "Digital History and the Politics of Digitization." *Digital Scholarship in the Humanities* 38, no. 2 (2023): 830–851. https://doi.org/10.1093/llc/fqac050.

Zimmermann, Patricia R. *Reel Families: A Social History of Amateur Film*. Bloomington: Indiana University Press, 1995.