Between History and Poetics: Identifying Temporal Dynamics in Large Audiovisual Collections

Nanne van Noord †
Mila Oiva *
Isadora Campregher Paiva †

Daniel Chávez Heras ‡
Carlo Bretti †
Ksenia Mukhina *
Tillmann Ohm *

Abstract

In this paper, we present a composite-methods approach to identify local and global temporal structures in large audiovisual collections.

Existing computational methods in the study of visual culture have been useful in identifying visual patterns across large collections of images [1, 2]. However, they have also been shown to be very limited when it comes to analysing the subjective experience of specific images by specific audiences [3, 4]. In the case of moving images, these limitations compound with a lack of attention to temporal form, which plays a key role in the production and interpretation of moving images, but has so far been largely ignored by computational methods.

Drawing from moving image scholarship about time [5, 6], and under the general principle of aesthetically sensitive systems that are designed to access inter-subjective experience of media, our work aims to address these limitations by sensing and correlating two kinds of temporal structure: local (artefact) time, and global (collection) time.

Local structure corresponds to the notion of cinematic time, understood here as the subjective experience of time created in audiences through artifice and the manipulation of the film medium, such as editing techniques that lend rhythm and meaning to individual compositional units such as shots, scenes, clips and films. Boundaries between these units are often coded in technical conventions and thematised as moving image aesthetics.

Global structure corresponds to historical time, meaning in this context the subjective experience of an "era" as a relational category that is placed by audiences before or after other eras. Boundaries between eras are often marked by historical events like large-scale economic, social, and political change, armed conflict, or technology development.

As a case study, we use a sample from a collection of the Daily News newsreel series, produced by the Central Documentary Film Studios in Moscow between years 1944-1992. We apply one computational method for each type of temporal structure, and test our findings using history and aesthetic heuristics: shot scale classification [7] used to identify local relations, which we test against known medium-specific aesthetics (e.g. close ups of political figures); and CLIP embeddings [8] used to identify global collection relations, which we test against known themes in the reels and historical events (e.g. agricultural scenes after Perestroika).
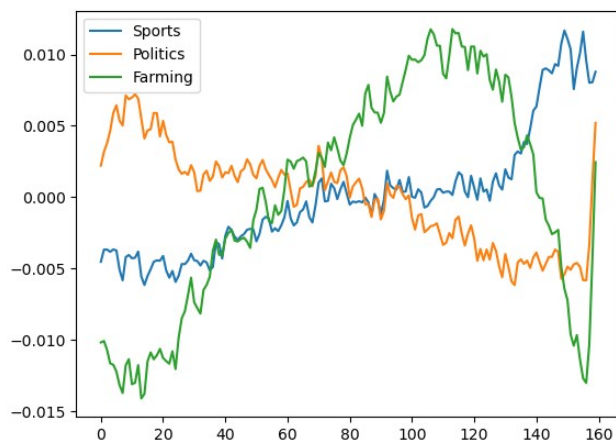
Fig 1. *Thematic blocs per reel over time.* Centered text-image similarities of queries on sports, politics, and farming average across the collection as measured by OpenCLIP. From this we can observe that first the newsreels are more related to politics, after which we see a spike for farming, and finally a peak for sports. This pattern of thematic blocs is consistent with initial assumptions drawn from sample viewings of the clips.

Through this approach we are able to identify editing patterns at the local level, and thematic patterns at the global level, and we are able to get a signal of the relation between the two temporal forms in these newsreels. Finally, we argue that this approach can be applied to other moving image archives and collections, signalling to a broader contribution to define the emerging field of computational moving image studies.

References

[1] L. Manovich, 'Visualization methods for media studies', in *Oxford Handbook of Sound and Image in Digital Media*, C. Vernallis, A. Herzog, and J. Richardson, Eds., Oxford: Oxford University Press, 2014, pp. 253–78.

[2] L. Manovich, *Cultural Analytics*. Cambridge, MA, USA: MIT Press, 2020.

[3] C. G. Olesen, 'Towards a "Humanistic Cinemetrics"?', in *The Datafied Society*, M. T. Schäfer and K. van Es, Eds., in Studying Culture through Data. Amsterdam University Press, 2017, pp. 39–54. doi: 10.2307/j.ctt1v2xsqn.7

[4] N. van Noord, M. Wevers, T. Blanke, J. Noordegraaf, and M. Worring, 'An Analytics of Culture: Modeling Subjectivity, Scalability, Contextuality, and Temporality'. arXiv, Nov. 14, 2022. doi: 10.48550/arXiv.2211.07460

[5] M. A. Doane, *The emergence of cinematic time: Modernity, contingency, the archive*. Harvard University Press, 2002.

[6] A. Somaini, E. Grignard, and M. Rebecchi, Eds., *Time Machine: Cinematic Temporalities*, 1st edition. Milano: Skira Editore, 2020.

[7] M. Savardi, A. Signoroni, P. Migliorati, and S. Benini, 'Shot scale analysis in movies by convolutional neural networks', in *2018 25th IEEE international conference on image processing (ICIP)*, Oct. 2018, pp. 2620–2624. doi: 10.1109/ICIP.2018.8451474

[8] G. Ilharco *et al.*, 'OpenCLIP'. Zenodo, Jul. 2021. doi: 10.5281/zenodo.5143773

Affiliations

† University of Amsterdam
* Tallinn University
‡ King's College London