

Kinetik und Methodik

Film als dynamische und multimodale Herausforderung für die DH

Bateman, John

bateman@uni-bremen.de
Universität Bremen

Diecke, Josephine

diecke@staff.uni-marburg.de
Universität Marburg

Ewerth, Ralph

ralph.ewerth@tib.eu
Universität Hannover

Heftberger, Adelheid

adelheidh@gmail.com
Bundesarchiv

Howanitz, Gernot

gernot.howanitz@uibk.ac.at
Universität Innsbruck, Austria

Spiegel, Simon

simon@simifilm.ch
Universität Zürich

Loertscher, Miriam

miriam.loertscher@zhdk.ch
Zürcher Hochschule der Künste

Motivation

Wo sind sie, die ‚kinetischen‘ Methoden für die Filmanalyse, die dem Film als sowohl dynamisches als auch multimodales Medium zumindest ein Stück weit gerecht werden können? Das geplante Panel versucht, der Beantwortung dieser Frage ein Stück näher zu kommen und adressiert damit ein Forschungsdesideratum in den DH: Filme sind – gerade auch aus quantitativer Sicht – notorisch schwer zu bändigen. Einerseits sind sie – multimodal – aus Bild, Text und Ton zusammengesetzt, andererseits bestehen sie – dynamisch – nicht nur aus Bildern, sondern eben aus *bewegten* Bildern, eine Tatsache, die gleich etliche Ebenen an Komplexität hinzufügt.

Verständlicherweise beschäftigen sich viele DH-Projekte gegenwärtig mit Einzelaspekten des Mediums Film, wobei hier in der Regel besonderes Augenmerk auf das Visuelle gelegt wird. Beispielsweise bietet das von Taylor Arnold und Lauren Tilton initiierte *Distant Viewing Toolkit*, ein etabliertes Python-Paket zur quantitativen Filmanalyse, zwar rudimentäre Aggregatoren für

Audio und Untertitel an, stellt aber ungleich mehr Möglichkeiten für die Analyse von Einzelframes zur Verfügung (Arnold / Tilton 2019). Diese Fokussierung auf einige wenige Teilelemente des Mediums ‚Film‘ wird häufig mit der im Vergleich zu Text oder Bild vielfach höheren Datenmenge argumentiert; ein Argument, das jedoch letztlich ins Leere laufen muss, können die DH ja gerade hier ihren Trumpf ausspielen.

Immerhin zeigen einige Forschungsprojekte aus der Computer Vision, dass neuronale Netze auch auf das Erkennen komplexer Bewegungsabläufe und letztlich ganzer Szenen bzw. szenischer Handlungen trainiert werden können und konkret bei der Erkennung von Gewaltszenen (Mumtaz et al. 2020) oder bei der automatisierten textuellen Beschreibung von Handlungsabläufen (Park et al. 2018) eingesetzt werden. Diese innovativen Methoden aus dem Deep Learning machen es noch dringlicher, gesamtheitlichere Zugänge zum Medium Film zu finden, und dieses nicht einfach als ‚bag of images‘ misszuverstehen, weil sie aufeinander abgestimmt und in den methodischen Kontext der Geisteswissenschaften eingebettet werden – ein Ansinnen, das uns vor Herausforderungen stellt. So nimmt es nicht wunder, dass die DH nicht die neuesten Methoden der Computer Vision einsetzt (vgl. dazu die Methodenübersicht in Pustu-Iren et al.).

Nun gibt es in den letzten Jahren auch im deutschsprachigen Raum einige große Forschungsprojekte, die eine ganzheitlichere Betrachtung des Mediums Film propagieren. So haben etwa das FilmColors-Projekt mit dem Annotationstool VIAN (ERC Advanced Grant, Zürich; Flückiger / Halter 2020), der von der Volkswagen Stiftung geförderte „Digital Cinema Hub“ (Marburg, Frankfurt und Mainz) sowie „TIB AV Analytics“ (Hannover, DFG) erste vielversprechende Versuche geliefert und setzen gleichzeitig sowohl in methodologischer Sicht als auch hinsichtlich der Forschungsfragen verschiedene Schwerpunkte. Zu klären bleibt deshalb die Frage: Wie führt man Ergebnisse aus Deep Learning mit visuellen Medien, Audio und Text mit empirischen Ergebnissen, Eye-Tracking-Analysen, etc. zusammen?

Das geplante Panel möchte aber nicht einfach die eben skizzierten Probleme beschreiben, sondern Wissenschaftlerinnen und Wissenschaftler aus den unterschiedlichsten Disziplinen und Forschungskontexten zusammenbringen, um mögliche Lösungswege aufzuzeigen und eine dem Medium Film entsprechende ‚kinetische‘ Methodologie zu skizzieren. Dabei wird nicht nur eine (forschungsgeleitete) Anwendungsperspektive eingenommen; vielmehr sollen ebenso Aspekte aus der Lehre, der (Film-)Kulturvermittlung und der Theoriebildung berücksichtigt werden.

Organisation des Panels

Das Panel verschreibt sich ganz einer lebendigen Diskussionskultur. Aus diesem Grund verzichten wir auf Kurzreferate der Teilnehmerinnen und Teilnehmer und setzen auf pointierte Eröffnungsstatements. Dabei bringt jede Teilnehmerin und jeder Teilnehmer eine spezifische Perspektive, eigene Forschungsfragen und Anliegen in die Diskussion mit ein, die im folgenden Abschnitt kurz umrissen werden. Die starke Fokussierung auf das Gespräch soll auch helfen, das Publikum verstärkt anzusprechen. Um der Diskussion einen Rahmen zu geben, verschickt die Moderation im Vorfeld Leitfragen, die helfen sollen das Gespräch in Gang zu setzen, zu fokussieren und – falls notwendig – zu einem roten Faden zurückzukehren.

Der Zeitplan sieht eine fünfminütige Einführung durch den Moderator vor, die von fünf zweiminütigen Kurzstatements der Teilnehmerinnen und Teilnehmer abgerundet werden. Es folgt ein of-

fenes Panelgespräch entlang der vorher verschickten Leitfragen im Umfang von 45 Minuten, bei dem das Publikum natürlich auch gerne eingreifen darf, die letzten dreißig Minuten sind explizit für Fragen aus und Diskussion mit dem Publikum vorgesehen; dabei möchten wir auch mit einer Social-Media-Komponente (Twitter) und anderen mit etablierten digitalen Feedbacklösungen aus der Hochschuldidaktik experimentieren, die das Eintreten in einen Dialog möglichst niederschwellig gestalten sollen.

Spezifische Perspektiven

Prof. Dr. John Bateman (Bremen) widmet sich der Frage, inwieweit Analysetools für dynamische Medien auch dynamisch sein müssen. Sind dynamische Werkzeuge zur Auseinandersetzung mit Daten für die Theoretisierung geeignet oder sind sie eher Hilfsmittel zur Datenexploration in Abgrenzung zur Theoriebildung? Was könnten dynamische Theorien sein? Ist der Begriff in sich kohärent? Inwieweit könnten dynamische Werkzeuge für den Umgang mit Daten und Datenanalysen von unserem Wissen über die Funktionsweise von Film, insbesondere von kinematografischen Ausdruckstechniken, profitieren?

Dr. Adelheid Heftberger (Bundesarchiv Berlin) interessiert, wie es der DH-Community (das schließt nicht nur Wissenschaftler*innen, sondern auch Kulturerbe-Verantwortliche ein) gelingen könnte, von Pilotprojekten, ‚Hacks‘, Einzellösungen für Projekte und fertigen Softwarelösungen – sprich: Einzelforschungsinteressen – hin zu Anwendungen in großem Stil (Fernseharchive) zu kommen. Eine wesentliche Frage ist dabei: Welche Fragestellungen sind überhaupt möglich, wenn eben nicht mit professionellen Tools gearbeitet werden kann?

Josephine Diecke (Marburg) sieht ihren Schwerpunkt in der Reflexion von ‚kinetischen‘ Methoden für die Filmanalyse im Kontext der Lehre. Im Rahmen des VW-Projektes „Digital Cinema Hub“ an den Universitäten Marburg, Frankfurt und Mainz beschäftigt sie sich mit der Integration von digitalen Tools und Methoden in die Lehre und Forschung. Als zentrale Herausforderung für die Film- und Medienwissenschaft sieht sie die Verknüpfung von aktuellen Tools für die Filmanalyse mit etablierten Methoden bzw. auch die bewusste Abgrenzung dieser beiden Pole.

Prof. Dr. Ralph Ewerth (Hannover) weist auf die großen Fortschritte hin, die Deep Learning für die Mustererkennung von audiovisuellen Daten gebracht hat. Von besonderem Interesse ist, für welche Analysen (z.B. Szenengrenzen, Kamerabewegung, Spracherkennung) Methoden direkt anwendbar sind und welche neuen Möglichkeiten sich für die systematische Filmanalyse ergeben. Weiterhin ist eine offene Frage, wie solche Algorithmen auf einfache Weise für Forschende in den Filmwissenschaften verfügbar und anwendbar werden. Hierzu können Ergebnisse des DFG-Projekts „TIB AV Analytics“ diskutiert werden, im Rahmen dessen in interdisziplinärer Kooperation zwischen Filmwissenschaft und Informatik eine offene Web-basierte Analyse-Plattform für diese Zwecke entwickelt wird.

Dr. Miriam Loertscher und PD Dr. Simon Spiegel (Zürich) haben im Rahmen des ERC-Forschungsprojekts „FilmColors. Bridging the Gap Between Technology and Aesthetics“ (PI: Prof. Dr. Barbara Flückiger) zusammen mit anderen die Annotationssoftware VIAN entwickelt, die im Rahmen der Keynote der DHD 2019 vorgestellt wurde und mittlerweile um neue Funktionen erweitert worden ist. So kann VIAN nun auch die Daten von Eye-Tracking-Experimenten verarbeiten und anzeigen; zudem wird die Software aktuell mit Funktionen für die Gesprächsanalyse erweitert. Besonders spannend erscheint es im Kontext des Panels,

auf die Verbindung von ästhetischen Parametern und empirischen Daten einzugehen, die VIAN leisten soll.

Bibliographie

Arnold, Taylor / Tilton, Lauren (2019): „Distant Viewing: Analyzing Large Visual Corpora“, in: *DSH*, fqz013, <https://doi.org/10.1093/digitalsh/fqz013> [letzter Zugriff: 15. Juli 2021].

Burghardt, Manuel / Heftberger, Adelheid / Pause, Johannes / Walkowski, Niels-Oliver / Zeppelzauer, Matthias (2020): „Film and Video Analysis in the Digital Humanities – An Interdisciplinary Dialog“, in: *DHQ* 14.4 <http://digitalhumanities.org/dhq/vol/14/4/000532/000532.html> [letzter Zugriff: 15. Juli 2021].

Flückiger, Barbara / Halter, Gaudenz (2020): „Methods and Advanced Tools for the Analysis of Film Colors in Digital Humanities“, in: *DHQ* 14.4, digitalhumanities.org/dhq/vol/14/4/000500/000500.html [letzter Zugriff: 15. Juli 2021].

Mumtaz, Aqib / Bux Sargano, Allah / Habib, Zulfiqar (2020): „Fast Learning Through Deep Multi-Net CNN Model For Violence Recognition In Video Surveillance“, in: *The Computer Journal*, bxaa061, <https://doi.org/10.1093/comjnl/bxaa061> [letzter Zugriff: 15. Juli 2021].

Park, Jae Sung / Rohrbach, Marcus / Darrell, Trevor / Rohrbach, Anna (2018): „Adversarial Inference for Multi-Sentence Video Description“, in: *arXiv*, 1812.05634, <https://arxiv.org/abs/1812.05634>.

Pustu-Iren, Kader / Sittel, Julian / Mauer, Roman / Bulgakowa, Oksana / Ewerth, Ralph (2020): „Automated Visual Content Analysis for Film Studies: Current Status and Challenges“, in: *DHQ* 14.4, <http://digitalhumanities.org/dhq/vol/14/4/000518/000518.html> [letzter Zugriff: 15. Juli 2021].

Shou, Zheng / Lin, Xudong / Kalantidis, Yannis / Sevilla-Lara, Laura / Rohrbach, Marcus / Chang, Shih-Fu / Yan, Zhicheng (2019): „DMC-Net: Generating Discriminative Motion Cues for Fast Compressed Video Action Recognition“, in: *arXiv*, 1901.03460, <https://arxiv.org/abs/1901.03460> [letzter Zugriff: 15. Juli 2021].