

VAnnotatoR: Ein Werkzeug zur Annotation multimodaler Netzwerke in dreidimensionalen virtuellen Umgebungen

Abrami, Giuseppe

abrami@em.uni-frankfurt.de

Goethe Universität Frankfurt, Deutschland

Spiekermann, Christian

s2717197@stud.uni-frankfurt.de

Goethe Universität Frankfurt, Deutschland

Mehler, Alexander

mehler@em.uni-frankfurt.de

Goethe Universität Frankfurt, Deutschland

Abstract

Die Verarbeitung, Erzeugung und Visualisierung von multimodalen Netzwerken, in denen etwa symbolische und ikonographische Inhalte miteinander vernetzt werden, um sie multimodal rezipierbar zu machen, bilden neuartige Herausforderungen für die digitalen Geisteswissenschaften, da sie weit über textbasierte Medien hinausreichen. Digitalisate von Schrifttexten, Bildern und Illustrationen können in diesem Zusammenhang als Beispiele für knotenbildende Informationsobjekte ebenso genannt werden wie dreidimensionale Objektrepräsentationen. Die Anordnung dieser Objekte zu *semiotischen Netzwerken* hat den Zweck, deren Zusammenhangsstruktur zu visualisieren und rezipierbar zu machen. Der Begriff des Netzwerks erlaubt es dabei, Mikrostrukturen auf der Ebene von Knotenpaaren oder -triaden ebenso zu unterscheiden, wie Mesostrukturen von so genannten Netzwerkmotiven oder gar Makrostrukturen, welche die Gesamtorganisation des jeweiligen Informationsraums thematisieren (etwa im Sinne einer Zentrums-Peripherie-Struktur). Multimodale Netzwerke waren bereits Gegenstand einer Reihe von Projekten (siehe z.B. Li et al. 2007; Kane, Alavi 2008; Nasser, Gleich 2017). Allerdings befindet sich darunter kein Projekt, das eine generische Plattform zur Visualisierung und Annotation solcher Netzwerke im Rahmen virtueller Umgebungen zur Verfügung stellt. Um diese Lücke zu schließen, haben wir **VAnnotatoR** (Spiekermann, Abrami, Mehler 2018) entwickelt. **VAnnotatoR** ist ein Framework zur Erstellung und Visualisierung multimodaler Netzwerke in virtuellen

Umgebungen. Das System erlaubt die Bearbeitung und Visualisierung dieser Netzwerke als *multimodale Hypertexte* (Mehler et al. 2018) in der *virtuellen Realität* (VR) und der *augmentierten Realität* (AR). **VAnnotatoR** wurde mithilfe von Unity3D¹ für VR implementiert und unterstützt gängige Desktop-VR-Brillen² (HTC Vive und Oculus Rift). Die AR-Implementierung von **VAnnotatoR** für Smartphones basiert wiederum auf AR-Core³. Mittels **VAnnotatoR** sind Inhalte visualisierbar, suchbar, relationierbar, attribuierbar und weit darüber hinaus annotierbar. Die derzeitige Implementierung dieses Systems erlaubt die Modellierung von Netzwerken auf der Ebene so genannter Hypergraphen (Berge 1989) und hierarchischer Graphen, in denen Knoten vollständige Graphen enthalten können. Abbildung 1 illustriert ein solches Netzwerk unter anderem durch Rekurs auf Texte, Bilder, Photographien, dreidimensionale Modelle und ein begehbare Haus. Um eine natürliche Interaktion mit den Komponenten solcher Graphen zu ermöglichen, beinhaltet **VAnnotatoR** eine Gesten- und Bewegungssteuerung mit dazugehörigen VR-Controllern. Da klassische VR-Eingabegeräte die Gestaltungsfreiheit der Gestensteuerung limitieren, beinhaltet **VAnnotatoR** darüber hinaus eine Gestenerkennung unter Verwendung von Leap-Motion⁴. Gleichzeitig wurde eine Gestenerkennung basierend auf Datenhandschuhen⁵ für den **VAnnotatoR** entwickelt (Kühn 2018). Da komplexe Annotationsprozesse i.d.R. von mehreren Annotatoren durchgeführt werden müssen, ermöglicht **VAnnotatoR** zudem die simultane und kollaborative Bearbeitung von multimodalen Netzwerken. Die Annotationen werden in UIMA-Strukturen, unter Verwendung des *UIMA-Database Interface* (Abrami, Mehler 2018), abgespeichert und nutzbar gemacht.

Mit **VAnnotatoR** ist ein Annotationswerkzeug entstanden, welches die Bearbeitung und Visualisierung komplexer Netzwerke gesteuert ermöglicht. Der Beitrag demonstriert dies am Beispiel einer *Public History of the Holocaust*, in welcher, basierend auf dem Konzept der *Stolperwege* (Mehler et al. 2017), **VAnnotatoR** dazu genutzt wird, dreidimensionale Modelle nicht mehr existierender bzw. unzugänglicher Gebäude begehbare zu machen und zu dokumentieren. Auf diese Weise entsteht ein begehbare und erweiterbare Informationsraum, der insbesondere dazu dient, historische Prozesse zu dokumentieren. **VAnnotatoR** ist daher als ein Werkzeug zu betrachten, das verschiedene geisteswissenschaftliche Disziplinen verbindet. Inwiefern **VAnnotatoR** daher als interdisziplinäres Werkzeug der digitalen Geisteswissenschaften aufzufassen ist, soll Diskussionsgegenstand der Präsentation dieses Systems sein.



Abbildung 1: Ein Bild und ein Text wurden aus einem Browser extrahiert (linker Kubus). Beide wurden segmentiert und mit anderen Entitäten relationiert. Das Netzwerk zeigt in diesem Beispiel unter anderem Relationen zwischen einem 3D Modell eines Gebäudes (welches im extrahierten Text thematisiert wird), einer Position (Karte), einer Videoübertragung aus der realen Welt, einer Personenrepräsentation, einem Audio-Dokument und einem Wikidata-Eintrag. Subnetzwerke können außerdem zu Knoten aggregiert werden (visualisiert als Kubus), welche untereinander wiederum relationiert werden können. Auf diese Weise entstehen virtuelle Netzwerke von annotierten Szenen oder Situationen, die von anderen Personen weiterverarbeitet oder rezipiert werden können. Die Legende zeigt die der jeweiligen Modalität zugewiesene Farbe.

Fußnoten

1. <https://unity3d.com/de>
2. Getestet mit HTC Vive (<https://www.vive.com>) und Oculus Rift (<https://www.oculus.com/rift>)
3. <https://developers.google.com/ar/discover>
4. <https://www.leapmotion.com>
5. Hi5 VR Glove (<https://hi5vrglove.com>)

Bibliographie

Abrami, Giuseppe / Mehler, Alexander (2018): *A UIMA Database Interface for Managing NLP-related Text Annotations*, in: Proceedings of the 11th edition of the Language Resources and Evaluation Conference. Miyazaki / Japan

Berge, Claus (1989): *Hypergraphs: Combinatorics of Finite Sets*. North Holland, Amsterdam.

Kane, Gerald C. / Alavi, Maryam (2008): *Casting the net: A multimodal network perspective on user-system interactions*, in: Information Systems Research, 19(3), pages 253–272.

Kühn, Vincent Roy (2018): *A gesture-based interface to VR*. Bachelor's thesis, Goethe-Universität Frankfurt.

Li, Zhi-Chun / Huang, Hai-Jun / Lam, William HK. / Wong, Sze Chun (2007): *A model for evaluation of transport policies in multimodal networks with road and parking capacity constraints* in: Journal of Mathematical Modelling and Algorithms 6(2), pages: 239–257.

Mehler, Alexander / Abrami, Giuseppe / Bruendel, Steffen / Felder, Lisa / Ostertag, Thomas / Spiekermann, Christian (2017): *Stolperwege: An App for a digital public history of the holocaust* in: Proceedings of the 28th ACM Conference on Hypertext and Social Media. Pages 319–320. New York, NY, USA. ACM.

Mehler, Alexander / Abrami, Giuseppe / Spiekermann, Christian / Jostock, Matthias (2018): *VAnnotatoR: A framework for generating multimodal hypertexts*, in: Proceedings of the 29th ACM Conference on Hypertext and Social Media. New York, NY, USA. ACM.

Nassar, Huda / Gleich, David F. (2017): *Multimodal network alignment* in: CoRR, abs/1703.10511.

Spiekermann, Christian / Abrami, Giuseppe / Mehler, Alexander (2018): *VAnnotatoR: a gesture-driven annotation framework for linguistic and multimodal annotation*, in: Proceedings of the Annotation, Recognition and Evaluation of Actions (AREA) Workshop. Japan.