

Visual Analytics – ein neues Forschungsgebiet an der Technischen Informationsbibliothek (TIB)

Ziel sind innovative Dienstleistungen für die Informationssuche in großen Medien- und Forschungsdatenbeständen

Ralph Ewerth

Im November 2015 wurde die Forschungsgruppe Visual Analytics in Verbindung mit der ersten Professur an der Technischen Informationsbibliothek (TIB) etabliert. Hiermit wird das Ziel verfolgt, die Forschungsaktivitäten an der TIB zu intensivieren und aus praxisrelevanten Ergebnissen mittelfristig innovative Dienstleistungen abzuleiten. In diesem Zusammenhang wurde das Themenfeld Visual Analytics identifiziert, das von besonderer Bedeutung für eine erfolgreiche Suche nach Informationen in großen Medien- und Forschungsdatenbeständen ist. Der umfangreiche und heterogene Datenbestand der TIB und die Bedürfnisse der Nutzerinnen und Nutzer ergeben eine Vielzahl von anwendungsbezogenen Forschungsfragen. Die Forschungsgruppe Visual Analytics kooperiert eng mit dem ebenfalls an der TIB angesiedelten Kompetenzzentrum für nicht-textuelle Materialien (KNM), das unter anderem das AV-Portal der TIB (<https://av.tib.eu>) entwickelt hat.

Was ist Visual Analytics?

Die Forschungsrichtung Visual Analytics ist relativ jung, der Begriff wurde vor etwas mehr als zehn Jahren geprägt. Keim et al. (2008) definieren Visual Analytics als die Kombination von automatischen Analyseverfahren und interaktiven Visualisierungsformen mit dem Ziel, (sehr) große Datenmengen für Nutzerinnen und Nutzer verstehbar und interpretierbar zu machen sowie diesen das Finden von Schlussfolgerungen und Entscheidungen auf Basis der Datenlage zu ermöglichen.

Visual-Analytics-Methoden und -Algorithmen sollen Nutzerinnen und Nutzer bei einer spezifischen Rechercheaufgabe unterstützen, um in dem interaktiven Suchprozess Informationen und Wissen aus (sehr) großen Datenmengen gewinnen zu können. In diesem Kontext spielen drei Komponenten beziehungsweise Prozesse eine wesentliche Rolle: 1.) die automatische Datenverarbeitung und Transformation von Da-

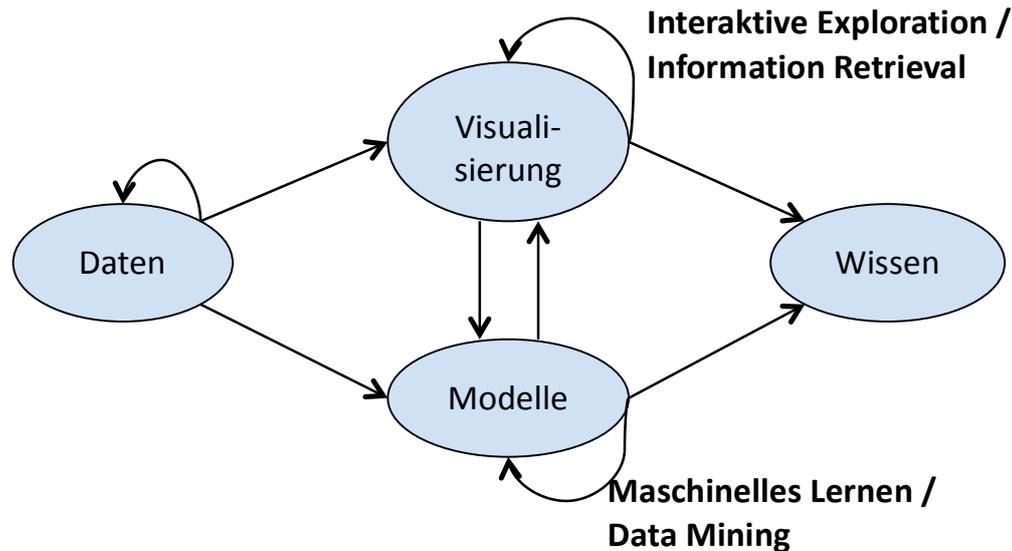
Die TIB adressiert mit einer Universitätsprofessur zu Visual Analytics die Erforschung von visuellen Analyse-, Such- und Präsentationsverfahren. In diesem Beitrag werden die Forschungsgruppe Visual Analytics der TIB sowie deren Zielstellungen vorgestellt. Zunächst wird kurz die Forschungsrichtung Visual Analytics erläutert und inwiefern diese für Dienstleistungen von digitalen Bibliotheken relevant ist. Im Anschluss werden exemplarisch zwei der Anwendungsfelder beziehungsweise Dienstleistungen präsentiert, die unter anderem im Fokus der Forschungsgruppe stehen werden: 1.) semantische Suche in Videodaten sowie 2.) explorative Suche in wissenschaftlichen Bilddatensammlungen beziehungsweise Abbildungen in Textdokumenten.

With a university professorship in Visual Analytics, TIB addresses research into visual analysis, search and presentation methods. TIB's Visual Analytics Research Group and its objectives are presented in this paper. First of all, the line of research in Visual Analytics will be briefly explained as well as the extent to which it is relevant for the services provided by digital libraries. Following this, two fields of application or services will be presented as examples of the aspects that the research group will focus on: 1.) semantic search in video data and 2.) explorative search in collections of scientific images or figures in text documents.

ten in eine für die Analyse und Exploration günstigere (nicht-visuelle) Repräsentation, 2.) die Generierung von neuen visuellen Repräsentationen oder Visualisierungsformen sowie 3.) die interaktive Exploration eines (großen) Datenbestandes auf Basis der beiden erstgenannten Punkte. Abbildung 1 illustriert den gesamten Prozess.

Oftmals wird Visual Analytics mit Informationsvisualisierung gleichgesetzt oder darauf reduziert, was jedoch die für einen analytischen Such- oder Rechercheprozess wesentlichen Punkte 1 und 3 außer Acht lässt. Aus Sicht der Forschungsgruppe sind an der TIB sowohl Datenbestände klassischer Art (Forschungsaufsätze, Patente etc.) als auch anderer und teils neuartiger Art wie etwa Video Abstracts, wissenschaftliche Filme, Vorlesungsvideos, 3D-Datenbanken, soziale Medien (zum Beispiel „Tweets“), 3D-Datenbanken, soziale Medien (zum Beispiel „Tweets“), Forschungsdaten sowie wissenschaftliche Software von Interesse.

Abbildung 1:
Zusammenhang
von Daten,
Data Mining
und interaktiver
Exploration im
Visual-Analytics-
Prozess, nach
Keim et al.
2008.



Forschungsthemen an der TIB

Die Forschungsgruppe adressiert neben dem Gebiet Visual Analytics auch die Bereiche Multimedia Information Retrieval sowie Usability-Fragen. Zu den Forschungsschwerpunkten zählen unter anderem:

- Automatische Annotation von visuellen Daten (Bilder, Video, 3D etc.)
- Informationsextraktion aus nicht-textuellen Daten
- Adaptive Klassifikationsverfahren für Daten aus verschiedenen Domänen
- Deep Learning zur Analyse und Annotation von Medien- und Forschungsdaten
- Ähnlichkeitssuche („query by example“)
- Intelligente Repräsentation und Visualisierung von großen Medienbeständen
- Interaktive Suche und Exploration von großen Medienbeständen
- Usability-Aspekte im Kontext der Suche in großen Medienbeständen

Der Begriff Medienbestände bezieht sich hier auf alle oben genannten Daten- und Medientypen. Die Forschungsgruppe beschäftigt sich prinzipiell sowohl mit textuellen als auch mit nicht-textuellen Medientypen (Bild, Video, Audio, 3D-Daten, Software etc.). In den folgenden Abschnitten illustrieren zwei Anwendungsbeispiele, wie angewandte Forschung Bibliotheksdienstleistungen erweitern beziehungsweise verbessern kann.

1. Beispiel: AV-Portal der TIB und andere Videoarchive

Das AV-Portal der TIB ermöglicht den Zugriff auf wissenschaftliche Filme und Videos mithilfe moderner Methoden zur automatischen Erschließung. Das

Portal wurde seitens der TIB in Zusammenarbeit mit dem Hasso-Plattner-Institut der Universität Potsdam entwickelt (siehe zum Beispiel Plank/Neumann 2014 oder Sack/Plank 2014).

Der grundlegende Prozess der Videoverarbeitung lässt sich wie folgt skizzieren: Wenn ein Video in das AV-Portal eingefügt wird, wird eine Reihe von automatischen Analyseverfahren durchgeführt. Hierzu gehört zunächst die zeitliche Segmentierung („Schnittterkennung“), die ein Video in dessen einzelne Einstellungen zerlegt. Zu diesen Einstellungen werden nun automatisch neue Metadaten extrahiert, welche die audiovisuellen Inhalte beschreiben und somit durchsuchbar und auffindbar machen. Zu den Analysealgorithmen gehören die Spracherkennung, die Erkennung von eingblendeten Texten mittels Video OCR (Optical Character Recognition) sowie die Konzeptdetektion zum Zwecke der automatischen Annotation der einzelnen Kameraeinstellungen (Hentschel et al. 2013). Unter der Konzeptdetektion versteht man die Erkennung von beliebigen audiovisuellen Inhalten, hierzu können beliebige Objekte, Ereignisse, Personen, Orte, Szenen etc. gehören.

Allerdings stammen die Videos des AV-Portals aus verschiedenen Domänen, welche zu einem Großteil den Fächern der TIB entsprechen: Technik sowie Architektur, Chemie, Informatik, Mathematik und Physik. Es ist offensichtlich, dass in der Domäne Architektur andere Konzepte relevant sind (zum Beispiel Konzepte wie Gebäudeart) als in Chemie-Videos (Konzepte wie Experiment, Strukturformel etc.). Zudem können sich audiovisuelle Konzepte innerhalb einer Domäne stark in ihrer Erscheinungsform unterscheiden oder sich mit der Zeit verändern. Diese Aspekte führen dazu, dass nicht immer optimale Ergebnisse für die auto-

matische Konzeptannotation erzielt werden können. Motiviert durch die genannte Problemstellung wird daher die Forschungsfrage verfolgt, wie sich Klassifikationsmodelle mit einem minimalen manuellen – oder bestenfalls ohne – Zusatzaufwand an neue Daten anpassen lassen. Hierzu werden seitens der Forschungsgruppe Verfahren aus dem Bereich maschinelles Lernen wie etwa Transferlernen oder transduktives Lernen (Ewerth et al. 2012a) beziehungsweise Möglichkeiten des domänenübergreifenden Lernens erforscht (Mühling et al. 2015). Eine weitere Forschungsfrage ist, wie im WWW verfügbare und bereits annotierte Daten für das Lernen visueller Konzepte optimal nutzbar gemacht werden können (Ewerth et al. 2012b).

Ein besonderer Fokus liegt gegenwärtig auf der Evaluierung und Erforschung von neuronalen Netzen und Deep-Learning-Methoden. Neuronale Netze sind mathematische Modelle, die sich an der Funktionsweise des menschlichen Gehirns orientieren, zum Lernen bestimmter (mathematischer) Funktionen. Tiefe neuronale Netze (Deep Neural Networks, Deep Learning) sind eine Variante, die sich unter anderem durch ihre große Anzahl von Schichten von herkömmlichen Netzen unterscheiden. Sogenannte Deep Convolutional Neural Networks haben unter anderem in der Bildanalyse hervorragende Ergebnisse erzielen können (zum Beispiel Krizhevsky et al. 2012). Durch die Verbesserung der Rechenkapazität moderner Rechner und Grafikkarten ist es möglich geworden, große Mengen von Lerndaten zu verarbeiten und somit Funktionen zur Erkennung komplexer Muster zu modellieren. In einigen Aufgabenstellungen zur Mustererkennung in Bildern wie etwa der Gesichtserkennung haben tiefe neuronale Netze bereits menschliche Leistungen erreicht (Taigman et al. 2014) beziehungsweise gar übertroffen (Schroff et al. 2015).

Mithilfe solcher Methoden wurden in einem kürzlich abgeschlossenen und von der Deutschen Forschungsgemeinschaft (DFG) geförderten Kooperationsprojekt mit dem Deutschen Rundfunkarchiv (DRA) und der Philipps-Universität Marburg 3.000 Stunden historisches TV-Material aus der ehemaligen Deutschen Demokratischen Republik (DDR) erschlossen. Das Material umfasste unter anderem Nachrichtensendungen wie etwa die „Aktuelle Kamera“. Die eingesetzten Analysealgorithmen umfassten eine zeitliche Video-segmentierung, Video OCR, audiovisuelle Konzeptdetektion inklusive DDR-spezifischer Konzepte sowie eine semantische Ähnlichkeitssuche (Mühling et al. 2016). Mittels der Deep-Learning-Verfahren konnten sehr gute Ergebnisse erzielt werden, so dass nunmehr eine feingranulare und qualitativ hochwertige



Prof. Dr. Ralph Ewerth

Der erste Professor an der TIB in Hannover leitet dort die Forschungsgruppe Visual Analytics

Prof. Dr. Ralph Ewerth, geboren 1972, hat seit November 2015 eine Professur am Institut für Verteilte Systeme (<https://www2.kbs.uni-hannover.de/ewerth.html>), der Fakultät für Elektrotechnik und Informatik an der Leibniz Universität Hannover und leitet die Forschungsgruppe „Visual Analytics“ (<https://www.tib.eu/de/forschung-entwicklung/visual-analytics/>) an der Technischen Informationsbibliothek (TIB). Dr. Ewerth studierte Informatik mit Nebenfach Psychologie in Frankfurt am Main und Marburg. Nach seiner Promotion in Marburg (2008) zum Thema „Robust video content analysis via transductive learning methods“ war er von 2008 bis 2012 verantwortlich für den Forschungsbereich Multimedia Computing am Lehrstuhl Verteilte Systeme der Philipps-Universität Marburg. Von 2012 bis 2015 war er Professor für Digitale Bildverarbeitung und Medientechnik an der Ernst-Abbe-Hochschule Jena und dort seit Anfang des Jahres 2015 auch Prodekan des Fachbereichs Elektrotechnik und Informationstechnik.

Suche in diesem Teilbestand des DRA möglich ist. Aktuell erforscht die Arbeitsgruppe, wie solche tiefen neuronalen Netze mit einem minimalen zusätzlichen Trainingsaufwand hinsichtlich der manuellen Erstellung von passenden Trainingsdaten für einen neuen Bereich (Domäne) beziehungsweise domänenübergreifend nutzbar gemacht werden können, zum Beispiel unter Verwendung von im WWW zugreifbaren und annotierten Bildern (Springstein/Ewerth 2016). Mittelfristig sollen die Ergebnisse solcher Forschungsbemühungen in das AV-Portal integriert werden, um so die Qualität der Suche sukzessive zu verbessern.

Zudem sollen auf Basis dieser semantischen Annotationen neue Visualisierungs- und Explorationsformen erforscht werden.

2. Beispiel: Unterstützung bildwissenschaftlicher Forschung sowie Suche nach Abbildungen in wissenschaftlichen Texten

Bilder spielen in verschiedenen bildwissenschaftlichen Fragestellungen sowie in Publikationen eine besondere Rolle. Die allgemeine Problemstellung der Suche in Bilddatenbanken (Image Retrieval) wurde in den vergangenen 15 bis 20 Jahren intensiv erforscht. Einfachere Verfahren zur Suche in Bilddatenbanken basieren in der Regel auf sogenannten Low-level-Merkmalen, die in technischer Hinsicht Formen, Farben oder Texturen beschreiben. Allerdings entsprechen Ergebnisse, die auf solchen Merkmalen basieren, sehr oft nicht den Suchbedürfnissen der Nutzerinnen und Nutzer („semantic gap“), die meist inhaltlicher oder semantischer Natur sind (Smeulders et al. 2000). In den vergangenen Jahren sind erhebliche Fortschritte erzielt worden, visuelle Inhalte („Konzepte“) in Bildern automatisch zu erkennen (Objektbeziehungswise Konzeptdetektion), insbesondere durch Ansätze des Deep Learning (Krizhevsky et al. 2012). Aufgrund dieser Fortschritte können nunmehr Suchanfragen inhaltlicher Art besser beantwortet werden, da die Güte der automatischen Annotation zum Teil die Qualität menschlicher Annotationen erreicht (He et al. 2015).

Die Forschungsgruppe Visual Analytics beschäftigt sich mit Fragestellungen zur semantischen Suche in Bilddatenmengen, die entweder in bildwissenschaftlichen Fragestellungen oder in wissenschaftlichen Publikationen relevant sind. So ist etwa gegenwärtig gemeinsam mit Kunsthistorikern und Informatikern der Universität Paderborn ein Projekt geplant, das im Bereich Digital Humanities anzusiedeln ist. In diesem Projekt sollen spezifische Bildanalyseverfahren zur inhaltsbasierte Recherche in Bilddatenbanken, welche Abbildungen beziehungsweise Digitalisate von Kunstobjekten enthalten, entwickelt werden. Auch hier sind adaptive Lernverfahren relevant, wie sie für die Analyse von Videos verwendet werden.

Zum anderen wird ein weiterer Fokus der Forschungsaktivitäten auf die Verbesserung der Suchmöglichkeiten von Abbildungen und Diagrammen in Forschungsartikeln und -berichten sowie in Patenten abzielen. Die TIB ist einerseits Depotbibliothek für Forschungsberichte von Vorhaben, die von bestimmten Bundesministerien gefördert wurden. Zum anderen gewährt sie als real-virtuelles Patentinformations-

zentrum (PIZ) in Niedersachsen Zugriff auf eine große Anzahl von Patenten. Abbildungen und Diagramme in wissenschaftlichen Publikationen enthalten in der Regel wichtige Informationen. Hier sollen Methoden zur Informationsextraktion sowie zur Herstellung von Text-Bild-Bezügen erforscht werden und eine bessere Suche in den Beständen der TIB ermöglichen. Auch hier erscheint die Verwendung und Erforschung von Deep-Learning-Ansätzen aussichtsreich. So haben etwa Karpathy und Fei-Fei (2015) vielversprechende Ergebnisse hinsichtlich der textuellen Beschreibung des Bildinhalts von herkömmlichen Fotos publiziert.

Zusammenfassung

In einem kurzen Überblick wurden in diesem Beitrag die Aktivitäten der Forschungsgruppe Visual Analytics der TIB dargestellt. Hierzu wurden exemplarisch zwei Themenfelder etwas näher beleuchtet: Semantische Suche nach Videos mit dem Anwendungsfall AV-Portal der TIB sowie semantische Suche in großen Bilddatenmengen von wissenschaftlicher Relevanz. Die Forschungsaktivitäten beschränken sich aber nicht auf diese Medientypen, sondern umfassen auch die Verbesserung der Suche und Exploration von Datenbeständen anderer Medientypen (3D-Daten, Text, soziale Medien, Forschungsdaten, Software etc.). ■

Literatur

1. Ewerth, R., Mühling, M. und Freisleben, B.: Robust Video Content Analysis via Transductive Learning. In ACM Transactions on Intelligent Systems and Technology, Vol. 3, No. 3, Art. No. 41, 1-26, 2012a.
2. Ewerth, R., Ballafkir, K., Seiler, D., Mühling, M. und Freisleben, B.: Long-Term Web-Supervised Learning via Random Savannas for Concept Detection in Images. In IEEE Transactions on Multimedia, Volume 14 Issue 4, Part 1, 1008-1020, 2012b.
3. He, K., Zhang, X., Ren, S. und Sun, J.: Delving Deep into Rectifiers: Surpassing Human-level Performance on Imagenet Classification. In Proceedings of the IEEE International Conference on Computer Vision, 2015.
4. Hentschel, C., Blümel, I. und Sack, H.: Automatic Annotation of Scientific Video Material based on Visual Concept Detection. In Proceedings of 13th International Conference on Knowledge Management and Knowledge Technologies, Article No. 16, 2013
5. Karpathy, A. und Fei-Fei, L.: Deep Visual-semantic Alignments for Generating Image Descriptions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3128-3137, 2015.
6. Keim, D., Andrienko, G., Fekete, J. D., Görg, C., Kohlhammer, J. und Melançon, G. (2008). Visual Analytics: Definition, Process, and Challenges. In Information Visualization, Springer Berlin Heidelberg, 154-175, 2008.
7. Krizhevsky, A., Sutskever, I. und Hinton, G. E.: Imagenet Classification with Deep Convolutional Neural Networks. In Advances in Neural Information Processing Systems, 1097-1105, 2012.

8. Mühling, M., Meister, M., Korfhage, N., Wehling, J., Hörth, A., Ewerth, R. und Freisleben, B.: Content-Based Video Retrieval in Historical Collections of the German Broadcasting Archive. In Theory and Practice of Digital Libraries (TPDL), Hannover, zur Veröffentlichung angenommen, 2016.
9. Mühling, M., Ewerth, R. und Freisleben, B.: Improving Cross-Domain Concept Detection via Object-based Features. In Proceedings of 18th International Conference on Computer Analysis of Images and Patterns, Malta, Springer, 359-370, 2015.
10. Müller, E., Otto, C. und Ewerth, R.: Semi-supervised Identification of Rarely Appearing Persons in Video. In: Proceedings of ACM International Conference on Multimedia Retrieval (ICMR), New York, ACM, 381-384, 2016.
11. Plank, M. und Neumann, J.: TIB's Portal for Audiovisual Media: New Ways of Indexing and Retrieval. In: FLA Journal, Vol. 40, 17 – 23, 2014.
12. Sack, H. und Plank, M.: AV-Portal – The German National Library of Science and Technology's Semantic Video Portal, ERCIM News No. 96, 2014.
13. Schroff, F., Kalenichenko, D. und Philbin, J.: Facenet: A Unified Embedding for Face Recognition and Clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 815-823, 2015.
14. Smeulders, A. W., Worring, M., Santini, S., Gupta, A. und Jain, R.: Content-based Image Retrieval at the End of the Early Years. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1349-1380, 2000.
15. Springstein, M. und Ewerth, R.: On the Effects of Spam Filtering and Incremental Learning for Web-supervised Visual Concept Classification. In Proceedings of ACM International Conference on Multimedia Retrieval (ICMR), New York, ACM, 377-380, 2016.



Prof. Dr. Ralph Ewerth

Leiter der Forschungsgruppe
„Visual Analytics“
Technische Informationsbibliothek
(TIB)
Welfengarten 1 B
30167 Hannover
ralph.ewerth@tib.eu

Dietmar Dreier International Library Suppliers

Als internationaler Library Supplier ist Dietmar Dreier seit 1981 für europäische Bibliotheken erfolgreich tätig



Ihr Spezialist für:

- Wissenschaftliche Monographien
- E-Books und Datenbanken
- Fachbezogene Neuerscheinungsdienste (Print & E-Books)
- Graue und Antiquarische Literatur
- Shelf Ready Service

Ihr Partner für e-content:

- Verlagsunabhängige Beratung
- Vertrieb von ProQuest Ebook Central™ (ebrary & EBL)
- Vertrieb von Verlagsprodukten (Einzeltitel und Pakete)
- Dietmar Dreier E-Book User Guide (erscheint jährlich)

Unser E-Book-Portal – die ideale Unterstützung für Ihre Erwerbungsentscheidungen:

- Über 250 Verlagsmodelle online recherchierbar
- Direkter Zugriff auf die entsprechenden Titellisten
- Direkter Zugriff auf unsere aktuellen Angebote

