

3D-Scheibenbäume: Ein Visualisierungsparadigma zur Quantifizierung strukturierter Daten

Philipp Roßberger¹

¹Labor für Multimediale Systeme (MMLab),
Fachbereich Elektrotechnik und Informatik,
Hochschule für Angewandte Wissenschaften,
Berliner Tor 7, 20099 Hamburg, Deutschland
rossbe_p@informatik.haw-hamburg.de

Abstract: Der wachsende Umfang digitaler Informationssammlungen erfordert die Entwicklung neuartiger Visualisierungsmethoden, die dem Anwender eine einfache und intuitive Navigation in komplexen Datenmengen ermöglichen. In den letzten Jahren wurde eine Reihe von Arbeiten vorgestellt, die diesem Anspruch durch den Einsatz dreidimensionaler Darstellungsformen gerecht werden sollen. Diese Ausarbeitung beschreibt Ergebnisse einer Bachelorarbeit, die sich mit der Entwicklung eines neuen Visualisierungsparadigmas namens 3D-Scheibenbäume beschäftigt. Um den Nutzen dieser Darstellungstechnik zu testen, wurde eine Software entwickelt, die dynamisch dreidimensionale Scheibenbaumszenen im X3D- und VRML-Format erzeugt.

1 Einleitung

Die visuelle Datenaufbereitung ist ein wichtiges Hilfsmittel zur Entdeckung von Strukturen und Relationen in komplexen Informationsmengen. Bekannte Visualisierungstechniken wie Säulen- oder Balkendiagramme eignen sich z.B. zur zeitlichen Dokumentation von Werteschwankungen.

Die visuelle Informationsdichte eines zweidimensionalen Diagramms lässt sich auf unterschiedliche Weise steigern. Eine Möglichkeit besteht in der Nutzung einer weiteren räumlichen Achse, also einer dreidimensionalen Visualisierung. Die Anfang der 90er Jahre am Xerox Parc Institut entwickelten Visualisierungsmethoden setzen bewusst auf eine dreidimensionale Darstellung von Informationssammlungen. Zu den bekanntesten, damals entstanden Methoden zählen u.a. die Perspective Wall [MRC91] und die Cone Tree-Visualisierung von [RMC91].

Die hier beschriebene Bachelorarbeit leistet einen Beitrag zum Forschungsfeld der Informationsvisualisierung. Die dabei entwickelte Scheibenbaum-Visualisierungsmethode erlaubt die Analyse multivariater¹ Daten anhand quantifizierbarer Kriterien auf deren Basis dreidimensionale Scheibenbaum-Objekte erzeugt werden. Im Rahmen der Bachelorarbeit

¹von mehreren Variablen abhängigen

wurde als Datengrundlage eine lokale Kopie der Online-Enzyklopädie Wikipedia zur Erzeugung dreidimensionaler Scheibenbaumszenen benutzt. Die so generierten Szenen erlauben es dem Anwender u.a. das zeitliche Wachstum von Wikipedia-Artikelkategorien (z.B. Politiker, Künstler, etc.) miteinander zu vergleichen.

2 Verwandte Arbeiten

Eine Reihe jüngerer Beiträge aus dem Gebiet der Visualisierungsmethoden zur Analyse des zeitlichen und quantitativen Wachstums von Informationssammlungen dient als Grundlage der hier vorgestellten Arbeit.

Das „CandyTop“-Interface von [YMW03] ist eine dreidimensionale Visualisierung logischer und zeitlicher Zusammenhänge von Textdokumenten. Durch die Darstellung der Versionshistorie lassen sich die unterschiedlichen Entwicklungsstufen eines Dokumentes auf einen Blick erfassen. Der „Interactive Timeline Viewer (ItLv)“ von [MKFU02] dient einem ähnlichen Zweck wie das „CandyTop Interface“, stellt dabei aber die textuellen Unterschiede zwischen verschiedenen Versionen eines Dokumentes in den Vordergrund.

Die grösste visuelle Ähnlichkeit mit der in dieser Arbeit vorgestellten Scheibenbaum-Visualisierungsmethode hat das „CiteSpace“-Paradigma von [Che04]. Chen nutzt Verknüpfungslinien und Scheiben um Netzwerke von Co-Zitaten zu visualisieren. Als Datengrundlage für Chens Verfahren dienen bibliographische Datenbanken.

3 Das Scheibenbaumparadigma

Der Scheibenbaumalgorithmus generiert dreidimensionale Darstellungen verschiedenartiger Datenbestände (siehe Abb. 1). Die Grundidee hinter diesem Ansatz basiert auf der Morphologie von Bäumen, im speziellen Baumstämmen.

Dem biologischen Vorbild folgend, nutzt das Scheibenbaumparadigma Baumringe um Entwicklungsschritte von Dokumenten oder Informationssammlungen chronologisch und quantitativ zu visualisieren. Ein Baum repräsentiert eine Menge von Dokumenten, die in einem Kontext stehen. Bei der PediaVis-Software, die den Scheibenbaumalgorithmus implementiert und im folgenden Abschnitt beschrieben wird, werden die Artikelkategorien der Wikipedia durch Bäume repräsentiert.

4 Technische Umsetzung und Systemdesign

Die Erzeugung von Scheibenbäumen erfolgt mit Hilfe eines Algorithmus, der in Abhängigkeit von Artikeldatum, -umfang, -aufrufhäufigkeit und weiterer Parameter Objekte in eine 3D-Szene einfügt. Die PediaVis-Software implementiert diesen Algorithmus um den struk-

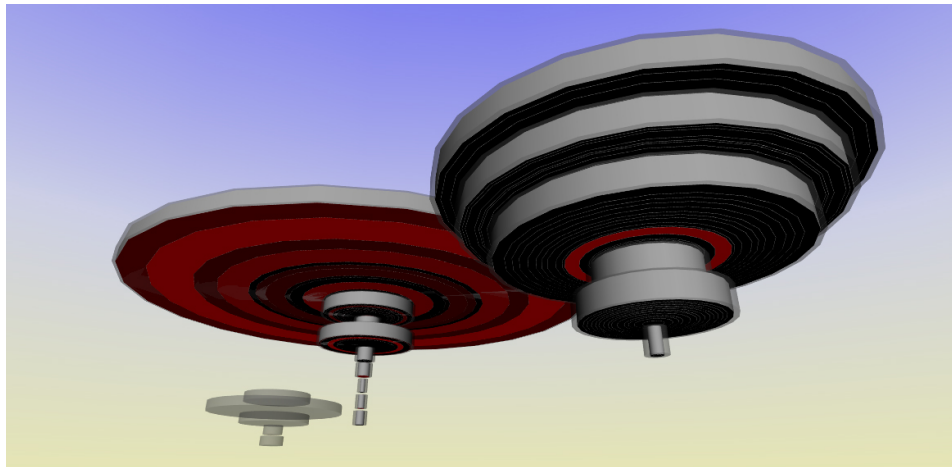


Abbildung 1: Scheibenbaumszene in der die zeitliche Entwicklung von drei Kategorien der Wikipedia gezeigt wird. Eine Baumscheibe umfasst alle Artikel, die innerhalb eines Monats zu einer Kategorie hinzugefügt wurden. Ein Baumring innerhalb einer Scheibe repräsentiert jeweils einen Artikel.

turellen Aufbau der Wikipedia-Enzyklopädie, aufgespalten in Kategorien und Artikel, zu visualisieren.

Die PediaVis-Software ist eine Java-Webapplikation, die in einer JSP-Engine (z.B. Tomcat) ausgeführt wird und deren Dienste vom Anwender durch eine GUI per Webbrowser genutzt werden. Die zentrale Funktionalität der Software besteht in der dynamischen Erzeugung von 3D-Scheibenbaumszenen und deren anschliessende Einbettung in HTML-Dokumente.

Der Anwender kann sich vor Erstellung und Darstellung einer Scheibenbaumszene entweder für X3D (Extensible 3D; eine XML-Modellierungssprache für 3D-Szenen) oder VRML als 3D-Datenformat entscheiden. Intern werden die 3D-Szene und darin enthaltene Objekte primär im X3D-Format erstellt und danach falls nötig per XSLT (Extensible Stylesheet Language Transformation) zu VRML umgewandelt. Die Struktur der Wikipedia-Kategorien und -artikel wird mithilfe einer lokalen Kopie der Wikipedia-Datenbank ermittelt, auf die von der Java-Applikation per SQL zugegriffen wird.

Die Dokumentation der schrittweisen Entwicklung der PediaVis-Software bildet einen Schwerpunkt der Bachelorarbeit. Zunächst wurden in der Analysephase funktionale und nicht-funktionale Anforderungen festgelegt um den Funktionsumfang der Software zu definieren. Auf Basis dieser Kriterien wurde in der anschliessenden Software-Designphase eine geeignete Systemarchitektur entwickelt. Hierbei wurden mithilfe von UML-Aktivitätsdiagrammen Kernkomponenten der Software identifiziert. Die daraus resultierende Systemarchitektur orientiert sich am MVC²-Entwurfsmuster.

Abschliessend wurden eine Reihe von Performanzmessungen durchgeführt, die das Lauf-

²Model-View-Controller

zeitverhalten der PediaVis-Software bei der Erzeugung von Scheibenbaumszenen mit unterschiedlich vielen Elementen zeigen.

5 Zusammenfassung

Ein Ziel dieser Arbeit war die Entwicklung eines neuartigen Visualisierungsparadigmas zur Quantifizierung strukturierter Daten. Ein Zielsetzung für weiterführende Arbeiten wäre die Anwendung des Scheibenbaumparadigmas auf eine andere Datengrundlage wie eines Dateiversionsverwaltungssystem (z.B. CVS) gut vorstellbar. Ferner wurde mit der PediaVis-Software eine Applikation zur dynamischen Generierung von Scheibenbaumszenen im VMRL- und X3D-Format realisiert. Die Entwicklung einer vielseitig einsetzbaren Produktionsplattform für diese 3D-Datenformate, ist ein interessantes Thema für zukünftige Arbeiten.

6 Danksagung

Diese Bachelorarbeit wurde unter der Betreuung von Prof. Dr. Renz und Prof. Dr. Kai von Luck an der HAW Hamburg angefertigt. Ich möchte mich an dieser Stelle besonders bei Prof. Renz für seine sehr hilfreichen Anregungen während der Entstehung meiner Arbeit bedanken.

Literatur

- [Che04] Chaomei Chen. *Information Visualization*. Springer, 2004.
- [MKFU02] Carlos Monroy, Rajiv Kochumman, Richard Furuta und Eduardo Urbina. Interactive Timeline Viewer (ItLv): A Tool to Visualize Variants Among Documents. In *Visual Interfaces to Digital Libraries [JCDL 2002 Workshop]*, Seiten 39–49, London, UK, 2002. Springer-Verlag.
- [MRC91] Jock D. Mackinlay, George G. Robertson und Stuart K. Card. The perspective wall: detail and context smoothly integrated. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems*, Seiten 173–176, New York, NY, USA, 1991. ACM Press.
- [RMC91] George G. Robertson, Jock D. Mackinlay und Stuart K. Card. Cone Trees: animated 3D visualizations of hierarchical information. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems*, Seiten 189–194, New York, NY, USA, 1991. ACM Press.
- [YMW03] Naofumi Yoshida, Jun Miyazaki und Akira Wakita. CandyTop Interface: A Visualization Method with Positive Attention for Growing Multimedia Documents. In *IV '03: Proceedings of the Seventh International Conference on Information Visualization*, Seite 360, Washington, DC, USA, 2003. IEEE Computer Society.