



**Hochschule für Technik
und Wirtschaft Berlin**

University of Applied Sciences

INDEPENDENT COURSEWORK BERICHT

Datanvisualisierung über WebVR

Caglar Özel

supervised by
Prof. Dr. Klaus Busch

Inhaltsverzeichnis

1	Abstrakt	1
2	Einführung	2
3	Virtual Reality	3
4	Daten Visualisierung	6
4.1	Der drang Daten zu Visualisieren	6
4.2	Arten und Techniken der Visualisierung	7
5	Server & Applikation	12
5.1	Installieren, Transpilieren des Client & Starten des Server	12
5.2	Server	14
5.3	Applikation	15
5.3.1	Anzeige & Konfiguration	15
5.3.2	Visualisierung der Daten	17
5.3.3	Interaktion mit der Welt	18
6	Fazit	19

Abstrakt

Virtual Reality ist eine Technologie die langsam aber sicher seinen Fuß auf dem Markt fasst. Anwendungsgebiete dieser Technologie sind noch sehr experimentell da die Software und Entwicklungsinfrastruktur noch nicht 100% ausgereift sind. Jedoch findet Virtual Reality eine weit größere Anwendung als nur in Spielen, Filmen, Technikdemos und 360 Grad Videos.

Einführung

Virtual Reality ist eine Technologie, welches von einer Fantasievorstellung oder einem Konzept aus Büchern, Filmen und Theorien, einen Weg in die praktische Umsetzung und damit seinen Weg in die Realität gefunden hat. Dies hat natürlich viel mit dem Fortschritt der Technologie in den letzten Jahrzehnten zu tun, welches mit der Steigerung der Leistung von Computern im Bereich der Rechenleistung, Speicherkapazität (RAM, Festplattenspeicher) und Grafikkartenleistung zusammen hängt. Zudem hat die einfache Verfügbarkeit dieser Technologie im Haushalt ebenbürtig seinen Einfluss auf diesen Fortschritt.

In diesem Report werden wir VR im Web verwenden um damit unbekannte Daten visuell darzustellen, desweiteren werden technologien, hürden und probleme angesprochen welches die Anwendung von VR im Web betreffen.

Virtual Reality

Virtual Reality mag zwar relative neu auf dem Markt sein, jedoch ist die Technologie seit Jahrzehnten in der Entwicklung und es gab mehrere Iterationen, welches verschiedene Versionen und Varianten der Technik für die virtuelle Welt hervorbrachte.

Das Konzept der virtuellen Welt wurde vermutlich zum ersten mal im Jahre 1935 vom Schriftsteller Stanley Weinbaum in der Science Fiction Story "Pygmalion's Spectacles" beschrieben.

Schon in dieser Geschichte verwendete der Hauptcharakter, eine Brille um in eine virtuelle Welt zu gelangen, wo seine Handlungen und Gefühle von und auf die reale Welt simuliert wurden. Dieses beschrieb relative akkurat welche Vorstellung und Visionen man hatte. [1]

Es folgten eine Reihe von Iterationen über die Jahrzehnte, welche den Werdegang der VR Brille von heute definierte.

Somit wurde im Jahre 1956 von Morton Heilig die "Sensorama" gebaut, welches man als die erste VR Maschine betrachten kann. Es bündelte mehrere Technologien um alle Sinne des Benutzers zu stimulieren. Für dieses Gerät wurden vier Filme gedreht, welches zu diesem Zeitpunkt als die Zukunft des Kinos betrachtet wurde. [1]



1960 veröffentlichte Heilig den "Telesphere Mask" welches das erste Head Mounted Display kurz HMD war. Es verfügte über die Funktion der Ausgabe von 3D Bildern und hatte eine Stereo Ausgabe. Dieses gerät verfügte jedoch noch nicht über die Funktion der Bewegungsverfolgung, welches mit dem Gerät "Headsight" von Ceomeau und Bryen zwei Philco Corporation Ingenieuren kam. Es verfügte über die Funktion der Bewegungsverfolgung des Kopfes, jedoch wurde dieses Gerät nicht als VR Brille verwendet sondern für das Militär entwickelt welches sie in riskanten Regionen als fern Steuerung von Kameras verwendeten. [1]

Ivan Sutherland welcher ein Informatiker in den Jahren 1965 war, veröffentlichte ein Paper namens Ultimate Display. In diesem beschrieb er das Computer Hardware die virtuelle Welt erstellen und in Echtzeit verwalten sollte. Sein Paper welches er veröffentlichte wird als der Bauplan vom heutigen VR Brillen gesehen. [1]

Von diesem Zeitpunkt an wurden die ersten HMD mit dem Fokus auf virtuelle Welten erfunden. Das gerät "The Sword of Damacles" welches im Jahre 1968 erschien wurde, trotz seiner Fähigkeit 3D Modelnetze abhängig von der Perspektive des Benutzers anzuzeigen, wegen seiner Größe und Notwendigkeit an einer Decke montiert zu sein nicht weiter als Labor Testphasen entwickelt. [1]



[2]

Schon zu diesem Zeitpunkt erkannte man das man VR für Trainingsimulationen und Medizinischen Behandlungen verwenden kann.

Somit wurde schon im Jahre 1979 von McDonnell-Douglas Corporation eine HMD für den militärischen Gebrauch entwickelt, dieses Gerät war in der lage die Augen des Benutzers verfolge um Bilder in Echtzeit passend zum Blickwinkel zu generieren. [1]



[3]

Oder im Jahre 1989 von der Nasa um Astronauten für anhand von VR auszubilden.



[4]

Einen Medizinischen gebrauch fand VR im Jahre 1997 durch Georgia Tech und Emory University welche den gebrauch von VR im Posttraumatischen Belastungsstörungen für Veteranen erforschten. Hier wurden Kriegsszenarios simuliert welches Virtual Vietnam benannt wurde um diese Symptome zu behandeln. Links zu Papern die zu dieser Behandlung veröffentlicht wurden. [1]

Virtual Vietnam Virtual Reality Exposure Therapy for PTSD

Wenn man wieder auf den kommerziellen Verkauf von VR Brillen zurück kommt. Waren Jaron Lanier und Thomas Zimmerman, gründer von VPL Research Inc., die ersten die VR Brillen, Handschuhe produzierten und für die Masse verkauften. Hier auf folgte bis auf den internen Gebrauch von VR Technologien wie von der Nasa oder medizinischen Experimenten bis zum Jahre 2010 nichts neues. [1]

Palmer Luckey, erzeugte den ersten Prototypen für die Oculus Rift welcher den Entwicklungsdrang von VR Technologien wieder neu entfachte. Folgen tut eine Kickstarter Kampanie im Jahre 2012 welche 2.4 Millionen USD sammelte und die Produktion der Oculus Rift Brillen in gang brachte.

Heutzutage hat jeder Hersteller (HTC, Sony, Apple, Google, Amazon, Samsung) seine eigenen VR Brillen und der Markt und Gebrauch von VR variiert von der Industrie, Medizien, Bildung, Unterhaltung bis hin zur Forschung. [1]

Daten Visualisierung

Da es sich in diesem Projekt hauptsächlich um die Visualisierung von Daten handelt, wird ein Teil dieses Dokuments den Aspekten und dem wissenschaftlichen Teil der Visualisierung von Daten gewidmet. Daten Visualisierung in der Informatik ist das Verwenden von grafischen Elementen um Zusammenhänge und Muster von Datensets zu offenbaren.

4.1 Der drang Daten zu Visualisieren

Menschen hatten und haben seit Jahrhunderten den drang Informationen visuell darzustellen und festzuhalten. Die Vorgeschichte der Visualisierung ist durch verschiedene Technologien und Bedürfnisse geprägt worden, diese variieren von Bildhauereien, Karten, Bildern bis hin zu Tabellen von Zahlen.

Doch die Frage lautet, warum Visualisieren wir eigentlich Daten?

Wie schon kurz erwähnt ist der Sinn der Daten Visualisierung, Zusammenhänge zwischen Informationen, Konzepten oder Logiken dem Menschen über eine Grafik zu zeigen und festzuhalten.

Diese können von mathematischen Prozessen, Zusammenhängen bis hin zu den Zusammenhängen von Ereignissen, Daten variieren. Es ist schwer Menschen Zusammenhänge anhand von Zahlen zu verdeutlichen ganz besonders wenn es mehr dimensionale Daten sind. Doch kann man diese Informationen aus einem gut durchdachten Grafen meist schneller, besser und einfacher extrahieren anstelle von Formeln oder Tabellen.

Dies hat mit unserem komplexen visuellem Kortex zu tun welches sehr gut darin ist Unterschiede in Form, Farbe und Größe zu ermitteln, wodurch es wiederum Zusammenhänge bei visueller Informationen sehr gut erfassen und verarbeiten kann. Der Bedarf Daten zu visualisieren stieg ganz besonders mit dem Anstieg der Daten die man Heutzutage im Internet finden kann.

4.2 Arten und Techniken der Visualisierung

Über die Jahrhunderte haben sich verschiedene Techniken und Arten der Datenvisualisierung durchgesetzt. In diesem Abschnitt werde ich etwas genauer auf Techniken und Arten von Visualisierungen eingehen.

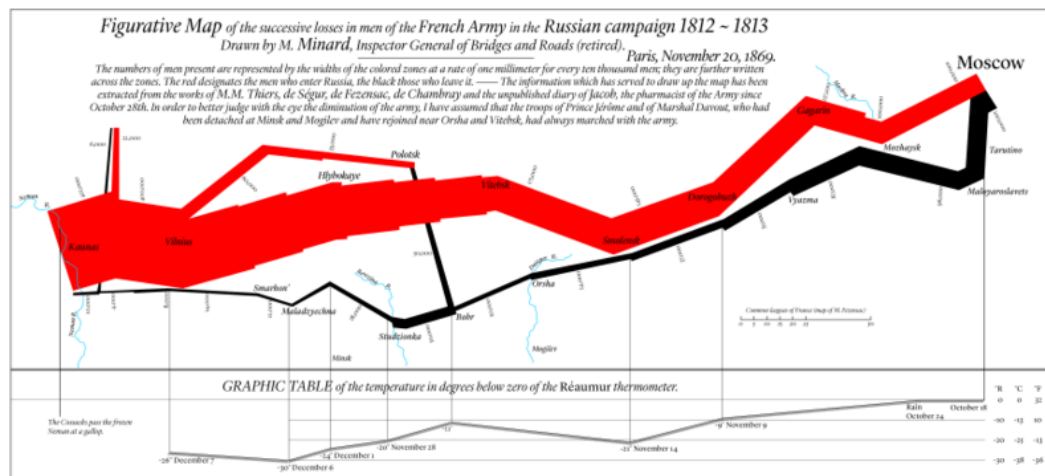
Es gibt verschiedene Arten von Visualisierungen welche verschiedene gut Daten visualisieren können, im folgenden werde ich eine Reihe von arten auflisten:

- Tabellen:

Als eine der ältesten Formen der Visualisierung welche wir im 21 Jahrhundert immernoch verwenden, ist die Tabelle eine gute Form der Visualisierung wenn man nur ein Gefühl für die Zahlen bekommen will. Jedoch ist schwer Anhand von Zahlen einen zusammenhang zu ermitteln da diese alleine stehend für sich nicht viel Aussagen.

- Charts: Balken oder Kreis:

Diese Form von Diagrammen sind üblicherweise zwei Dimensional, doch kann man in der Theorie n Dimensionale Grafiken erstellen. Zum beispiel Entwurf Charles Joseph Minard, ein französischer Bauingenieur, einen Grafen im Jahre 1869 welches den Marsch von Napoleon nach Russland beschrieb. [5]



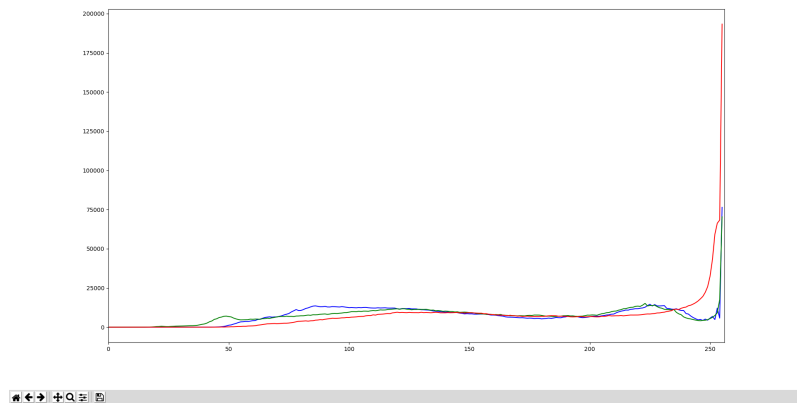
[6]

In diesem einen Grafen erfasst Minard folgende Informationen:

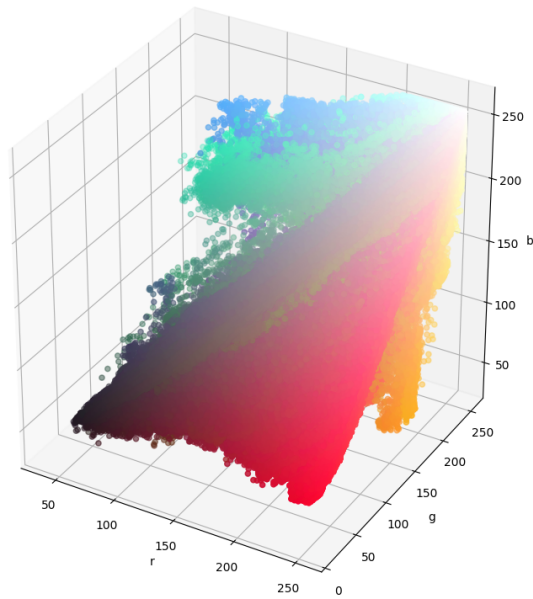
- Größe der Arme anhand der breite des Balken zu verschiedenen Zeiten
- Temperatur anhand des simplen Grafen unten
- Richtung des Marsches (Rot Hinweg, Schwarz: Rückweg)
- Ort anhand der Karte
- Distanz anhand der Karte
- Latitude, Longitude anhand der Karte

- Ein bis n dimensionale Diagramme:

Diese Sorte von Grafen sind sehr gut wenn man die Informationen in maximal drei Dimensionen, minimal einer Dimension hat oder die jeweilige andere durch Anwendung von Formeln oder anderen Gegebenheiten umwandeln kann. Zum Beispiel kann man die Farbverteilung eines Bildes sehr einfach und präzise in so einem zwei oder drei dimensional Raum als Diagramm visualisieren. Für das folgende Bild wird ein 2D Histogramm generiert welches die Menge der drei Grundfarben (rot, grün, blau) visualisiert.



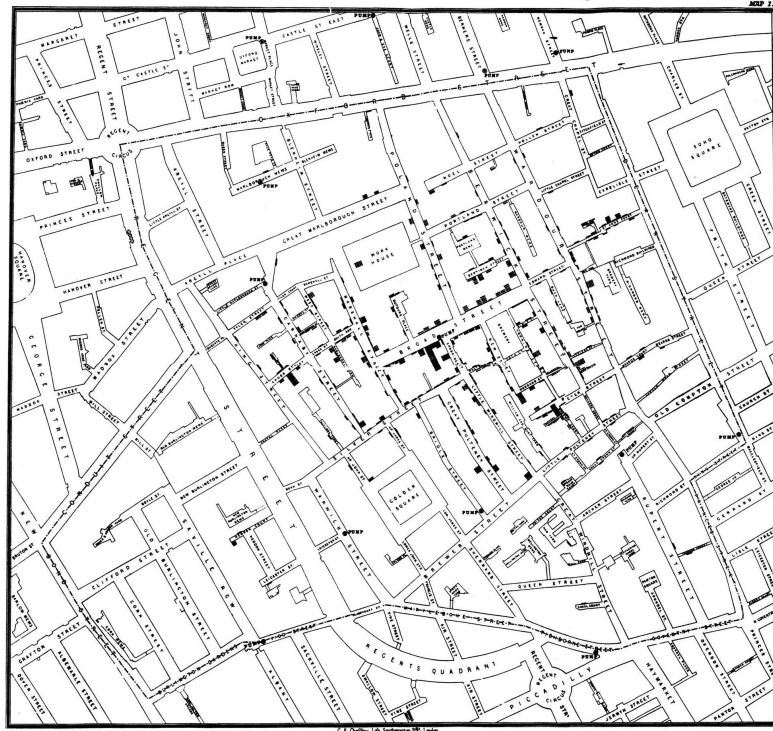
Wenn man für das selbe Bild nun ein 3D Histogramm generiert gewinnt man eine Dimension in der man eine weitere Information anzeigen kann. In diesem Beispiel bringt es den Information, wie die genaue Verteilung der drei Grundfarben in einem Raum aussieht.



azimuth=-60 deg, elevation=30 deg

- Karten:

Karten werden seit Jahrtausenden verwendet um Distanzen und Landzüge zu visualisieren, jedoch kann man mit einer Karte wie John Snow im Jahre 1854 bewies, Ursachen für ausbrüche von Cholera ermitteln. [5]

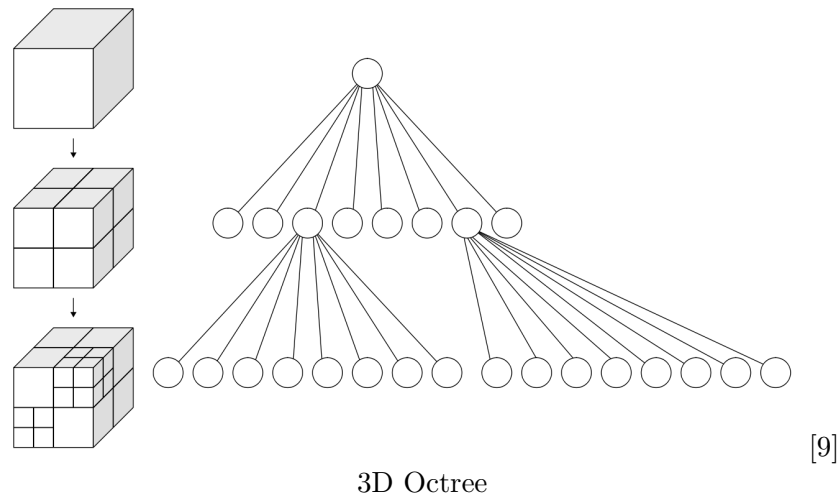
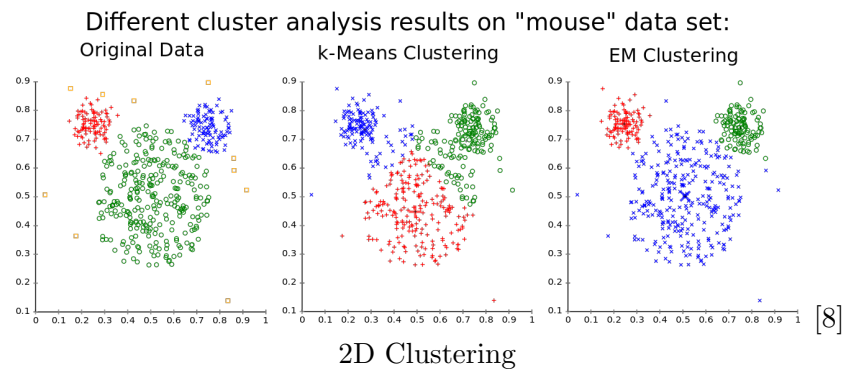


[7]

Auf dieser Grafik wurden Regionen mit Balken versehen, welches die Häufigkeit des Ausbruchs von Cholera indizierte. Des weiteren konnte dadurch ermittelt werden das die Ursache der hohen ausbruchrate in den Distrikten die Wasserquelle war, welches durch das Kanalsystem verpestet wurde. [5]

- 2D, 3D Raum:

Die Visualisierung in einem freien 2D oder 3D Raum ohne vordefinierte Axen, ist relative neu und fokussiert sich auf die verarbeiten und visualisieren von Daten, die eventuell keine direkten Zusammenhänge haben. Dieser drang kommt solche Daten zu visualisieren entsteht aus den Bereichen BigData und Machine Learning. Um diese Sorte von Daten zu visualisieren müssen sie erstmal Analysiert werden. Während der Analyse werden zusammenhänge in einer höheren oder niederen Ebene gefunden. Wenn diese zusammenhänge Erzeugt wurden kann man mit den Algorithmen k-means, octree oder anderen diese Daten in einem freien Raum visualisieren.



Server & Applikation

Nach der Analyse von verschiedenen Visualisierungsmethodiken musste der Umfang und die Komplexität der Applikation bestimmt werden. Folgende Kriterien wurden festgelegt:

- Visualisierung sollte unspezifisch auf Daten funktionieren
- Benutzer sollte in der Lage sein, die Struktur des Grafen selbst zu bestimmen
- 3D Welt sollte ohne weitere Benutzer Interaktion generiert werden

Des weiteren wurde festgelegt, das die Webseite mit dem Framework ReactJS, das zeichnen und erstellen der 3D Welt mit dem Framework AFrame und der Server der die Webseite veröffentlicht mit NodeJS geschrieben wird.

Die Applikation wird anhand des Package Managers "npm" aufgesetzt.

5.1 Installieren, Transpilieren des Client & Starten des Server

Da die Applikation über den PackageManager npm verwaltet und aufgesetzt wird, ist die Applikation namentlich in zwei Ordner unterteilt.

- Server
- Client

Diese beinhalten je eine tsconfig.json und package.json und im falle des Client eine webpack.config.js.

- package.json

Die package.json beinhaltet alle Abhängigkeiten des Clients und Servers. Man muss diese über den Befehl "npm install" sowohl für den Server als auch für den Client installieren.

- client: webpack.config.js

Webpack ist ein Modul welches mehrere JavaScript Dateien, in eine große JavaScript Datei bündelt. Der Zweck hiervon ist das man diese Einfach in einem Browser einbinden kann.

Um einen Prozess zu haben der Konsequent die Dateien überwacht (develop server), muss man den Befehl "npm start-watch" Ausführen. Dieser startet intern zwei Prozesse die Asynchrone laufen. Der Prozess "webpack -watch" welches die Typescript Dateien in JavaScript transpiliert und der Befehl "node-sass -watch ..." welches die Stylesheets kompiliert. Um nur einmal für einen Release zu bauen reicht der Befehl, "npm run start-build" aus.

- server starten

Der Server wie beschrieben ist ein NodeJS Server und kann unter dem Server Ordner mit dem Befehl, "npm run prod" gestartet werden. Man bedenke das vorher der Client mindestens einaml über den befehlt "npm run start-build" transpiliert werden muss, oder ein laufender build Server über den Befehl "npm start" existieren sollte.

Der Server holt sich momentan nur die Client Dateien und leitet diese zum Browser weiter. Mögliche und erstrebenswerte Erweiterungen sind die Verarbeitung von den Dateien auf der Server-seite (hebt das Limit der Datensatzgröße auf) und ein Anbindung an eine Datenbank welches die Daten beinhalten könnte.

- tsconfig.json

Die tsconfig.json beschreibt dem TSServer der die Typescript Dateien überwacht, nach welchen Standards sie geschrieben werden müssen.

5.2 Server

Die Entscheidung einen Server zu schreiben hat mehrere Gründe auf die ich in diesem Kapitel kurz eingehen werde.

Einer der Gründe ist das die Anwendung über einen Server veröffentlicht werden muss um im besten Fall diesen auf einen herkömmlichen Server im Internet hochladen zu können. Ein weiterer Grund war, das in dem Design Prozess die Idee existierte, einen MongoDB Server über Docker Instanz zu verwalten. Dieser würde die Datenstruktur, die Ausgewählt und Generiert wurde, abspeichern und zur abfrage jederzeit bereit halten.

Wie sich später auch herausstelle wäre der finale Grund den Server zu verwenden, das Browser einen lokalen Cache Limit haben. Somit können größere Dateien (500 MB und größer) nicht mehr direkt vom Browser verarbeitet werden. Diese müssten in einer produktiv Umgebung erstmal auf den Server hochgeladen. Dort analysiert, verarbeitet und umgewandelt werden, und zu guter Letzt wieder zu der Anwendung zurück geschickt werden.

5.3 Applikation

Basierend auf den Kriterien wurde die Funktionalität der Anwendung in drei Abschnitte unterteilt.

- Anzeige & Konfiguration
- Analyse
- Welt Generierung

Zu dem musste ich mir in diesem Stadium schon Gedanken machen wie meine Daten in der Praxis aussehen würden also wurde Datenquellen wie imdb angeschaut. Worauf hin fest gestellt wurde das der größte Teil von solchen Daten im CSV, TSV oder JSON Format ausgehändigt werden. Mit diesem Wissen ging ich an die Logik ran diese Daten anzuzeigen und zu konfigurieren.

5.3.1 Anzeige & Konfiguration

Die Anzeige & Konfiguration der Daten wird wiederum über drei Schritte absolviert. Zuerst muss eine valide csv, tsv oder json (json in einer flachen Hierarchie) ausgewählt werden.

Beispiel JSON

```
[
  {
    id: "0",
    title: "Geschichte des VR",
    author: "Walter Guenther",
    veroeffentlicht: "1996"
  },
  {
    id: "1",
    title: "Datenvisualisierung 101",
    author: "Peter Watson",
    veroeffentlicht: "2001"
  }
]
```

Falls eine CSV oder TSV Datei ausgewählt wurde, wird dieser intern verarbeitet und anhand eines D3 Moduls umgewandelt. Ein teil des Inhaltes wird dem Benutzer veranschaulicht. In diesem Beispiel verwenden wir den Datensatz von imdb und zeigen dem Benutzer die ersten 10 Einträge mit allen vorhandenen Attributen und ihren zugehörigen Werten an.

Select a File: title.basics_subset.tsv

Data Sample								
id	type	primaryTitle	originalTitle	isAdult	startYear	endYear	runtimeMinutes	genre
tt0011713	short	Snakeville's Hen Medic	Snakeville's Hen Medic	0	1915	\N	\N	Comedy,Short
tt0011714	movie	So Long Letty	So Long Letty	0	1920	\N	\N	Comedy
tt0011715	movie	Sol de gloria	Sol de gloria	0	1928	\N	\N	Drama,Romance
tt0011716	movie	Something to Think About	Something to Think About	0	1920	\N	78	Drama
tt0011717	movie	The Son of Tarzan	The Son of Tarzan	0	1920	\N	248	Action,Adventure
tt0011718	short	A Son of the North	A Son of the North	0	1920	\N	\N	Short,Western
tt0011719	short	The Song Birds	The Song Birds	0	1920	\N	\N	Animation,Comedy,Short
tt0011720	short	The Sour Violin	The Sour Violin	0	1920	\N	\N	Animation,Comedy,Short
tt0011721	movie	The Soul of Youth	The Soul of Youth	0	1920	\N	60	Drama
tt0011722	short	The Soilers	The Soilers	0	1923	\N	20	Comedy,Short

... Note that some Data may not be displayed if the Attributes are missing...

Hiernach wird dem Benutzer die Option angeboten, ein primären und sekundären Parameter auszuwählen. Des weiteren können weitere Werte definiert werden die im sekundären Knoten visualisiert werden.

Configuration

Primary Attribute:

genre

Secondary Attribute:

primaryTitle

Information visualized in the secondary Nodes:

id

type

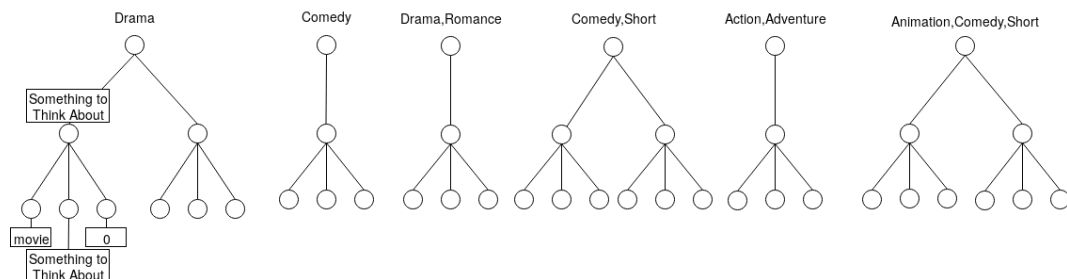
originalTitle

isAdult

Experimental options!

☐ Try getting Images

Diese Attribute werden verwendet und eine Parent, Child Hierarchie zu erstellen. Wenn man nur die oben ersharen Daten in Betracht zieht. Wird nach der Auswahl von den Attributen, eine interne Datenstruktur generiert die visuell Dargestellt wie Folgt aussieht.



5.3.2 Visualisierung der Daten

Mit dem Klick auf "Create VR World" werden die Daten an eine Scene weitergeleitet, womit anhand des Frameworks A-Frame eine 3D Welt generiert wird. Bei der Generierung dieser Welt musste Probleme wie Übersicht der Daten und Größe der Welt beachtet werden.

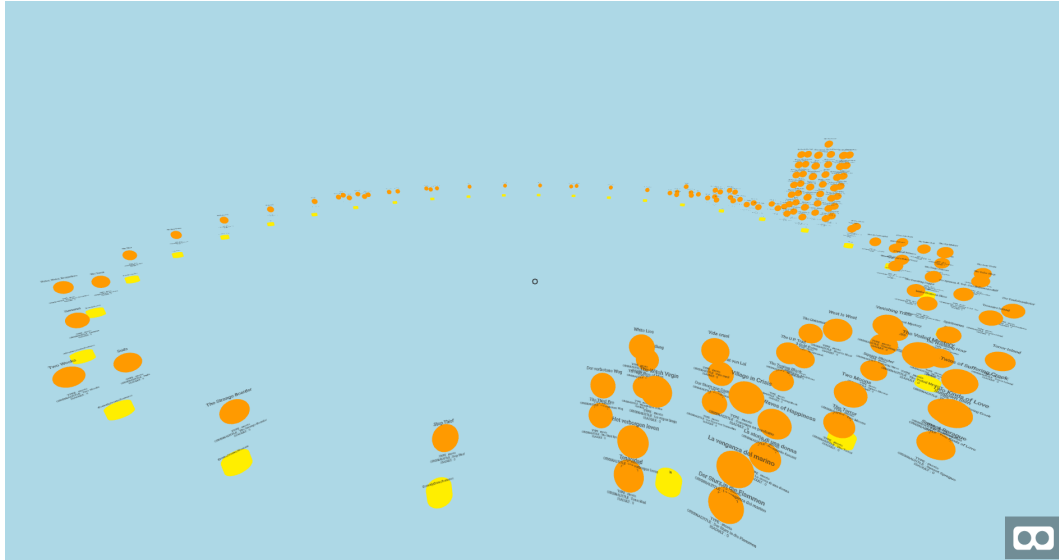
Wenn die Datensätze sehr groß sind, wird die Welt dementsprechend groß. Um dieses zu vermeiden wird der Datensatz vor der Welt Generierung in kleine Portionen von 30 Objekten unterteilt. Heißt wenn der Datensatz 200 Objekte beinhaltet wird dieses in sieben Teile unterteilt. Durch diesen Schritt wird die Performanz der Anwendung gesteigert und die Übersicht der Daten verbessert.

Eine zirkuläre Verteilung der Haupt- und Unterknoten wird in diesem Projekt als die Hauptmethode definiert. Um hierfür eine sinnvolle Verteilung der Daten zu erreichen, wird dynamisch für jeden Hauptknoten, ein Position in einem Radius abhängig von der Größe der Sektion ermittelt.

$$position = \begin{bmatrix} radius * Math.cos((index/length) * 2 * Math.PI) \\ y \\ radius * Math.sin((index/length) * 2 * Math.PI) \end{bmatrix}$$

Dies führt dazu das die Knoten, egal wie viele, gleichmäßig in einem Kreis verteilt sind. Die selbe Formel wird zur Berechnung der positionen der Unterknoten verwendet, jedoch werden diese relative zur Wurzel verschoben.

Bei den Unterknoten werden, um die Übersicht zu steigern, ebenfalls Untergruppen erzeugt und auf der y-Achse gestapelt. Gleichzeitig werden diese zuerst versteckt gezeichnet und müssen über die Interaktion mit dem Hauptknoten eingeblendet werden.



5.3.3 Interaktion mit der Welt

Für die Interaktion in der Applikation wurden eigene Events definiert und geschrieben. Um zum Beispiel unterknoten Einzublenden, muss auf den Hauptknoten geklickt oder per Controller X(Playstation) oder A(XBox) gedrückt werden. Man kann die Kamera im VR und Normal Modus per Controller oder Tastatur (wasd) steuern. Die Steuerung per Controller geschieht wie gewohnt mit dem linken Stick, wobei die Höhe der Kamera mit der vorderen rechten (senken) und linken (erhöhen) Schulter taste gesteuert wird. Das wechseln von Szenen / Sektionen geschieht über das Drücken des linken und rechten Analog tasten, links fürs nächstes- und rechts für vorherige Scene.

Fazit

Man sieht anhand dieser und anderer Applikation aus dem Internet das Virtual Reality für den Browser soweit ausgereift ist, das man komplexe und große 3D Welten gut und performant visualisieren kann. Natürlich müssen auch hier, wie in jedem Gebiet der Computergrafik und Software Entwicklung maßnahmen ergriffen werden um Performanz so hoch wie möglich zu halten.

Wenn es darum geht Daten mit dieser Technologie zu visualisieren sieht man, trotz Limitierung durch den Browser, das man ebenfalls mit Datenvisualisierungsansätzen gute Resultate erreichen kann.

Literaturverzeichnis

- [1] Dom Barnard. History of vr - timeline of events and tech development. <https://virtualspeech.com/blog/history-of-vr>.
- [2] Image sword of damacles. <https://virtualspeech.com/img/blog/header/history-of-vr.jpg>.
- [3] Image vital vr. <https://virtualspeech.com/img/blog/VITAL-helmet-vr.jpg>.
- [4] Image nasa vr. <https://virtualspeech.com/img/blog/virtual-environment-workstation-project-nasa.jpg>.
- [5] Data is beautiful: 10 of the best data visualization examples from history to today. <https://www.tableau.com/learn/articles/best-beautiful-data-visualization-examples>.
- [6] Image minard chart. https://upload.wikimedia.org/wikipedia/commons/thumb/e/e2/Minard_Update.png/800px-Minard_Update.png.
- [7] Image snow cholera map. https://cdns.tblsft.com/sites/default/files/pages/2_snow-cholera-map.jpg.
- [8] Image 2d clusters. https://upload.wikimedia.org/wikipedia/commons/thumb/0/09/ClusterAnalysis_Mouse.svg/1280px-ClusterAnalysis_Mouse.svg.png.
- [9] Image octree. <https://upload.wikimedia.org/wikipedia/commons/thumb/2/20/Octree2.svg/1280px-Octree2.svg.png>.