

Proposal for AniVision Data Visualization

Tobias Haider

Tuvshin Selenge

Christian Schönauer

Group 13

Introduction

This document describes our design of an interactive visualisation of the AniVision dataset. The first sections of the descriptions are structured following the main corners of the Data-User-Tasks Design Triangle. The last section describes how the visualisation prototype can be implemented addressing the different aspects of the Visualization Reference Model.

Data Description

The AniVision dataset contains data on the use of animation in educational, commercial, and news films from Austria, East and West Germany, focussing on the years 1945 to 1989 (but also containing videos from the 1930s and one as late as 2000). The provided data consists of two tables, namely videos and segments, which can be combined using the video_id key. While the videos table contains the descriptive information about a video, the segments table holds all the video effect information for all the videos.

Videos

We are predominantly looking into the following columns from the videos dataset:

- **Date of Production/Release:** Provides temporal information about when each video was produced and released, giving insights into the timeline of video content. (Unfortunately the available data is incomplete. Most often the production date is available. We are therefore using that and for videos, where only the release date is available, the release date is used. For the sake of visualisation it does not make much difference and usually there is only 1-2 years between production and release anyways. Videos without any date are ignored in the visualisation.)
- **Genre:** Specifies the type of content, such as "Science Film" or "Educational Film," allowing us to categorise and compare different videos.
- **Production Country:** Indicates the country responsible for the production of the video, helping us analyse geographical patterns in the dataset. During data exploration we found that 88% of the 225 videos have been produced in Western Germany. Therefore, we assume that it is very difficult to discover or visualize cultural exchange between countries of statistical significance.

- **Production Organization and Agents:** Provides details about the organisations and individuals involved as-well-as their roles in the production of the videos, giving context to the source.

Segments

In the segments dataset, we are focusing primarily on:

- **Image Type:** Describes the visual type of the segment (e.g., "Still Image," "Animation"), which helps us understand the nature of content presented in each part of the video.
- **Duration:** Defined by the start_time and end_time columns, indicating how long each segment lasts, allowing us to analyse the temporal structure of the videos.

Data Types and Measurement Levels

Nominal (Categorical):

- **Genre**
- **Production Organisation**
- **Agents** (list of individuals and their roles)
- **Image Type**

Ordinal (Discrete):

- **Production Country** (ranked by numeric codes)
- **Ratio** (continuous with a True Zero)

Quantitative:

- **Date of Production/Release** (Could also be considered ordinal when events are categorised into time periods)
- **Duration** (calculated from start_time and end_time)

Structural Analysis

Multivariate:

- Both tables in the AniVision dataset are multivariate, because they have multiple attributes, like Date, Genre, Production Country, Duration, Image Type etc.
- Furthermore the videos and segments tables are linked by the video_id column, increasing dimensionality in combined data tables
- Each video can have multiple segments, thus allowing analysis at different levels

Time-Oriented:

- The videos have a production and release date, thus making the data time-oriented.
- The segments dataset has start_time and end_time columns, indicating a time-oriented structure.

Tree/Hierarchy:

- The relationship between videos and segments resembles a tree/hierarchy structure.
- A video acts as the parent node, and its segments are the child nodes. This hierarchical relationship enables one to navigate from videos to their respective segments.

Users

The potential users which could be interested in exploring the data can be put in three different groups:

- Professionals (producers, cinematographers, editors)
- Academics (historians, students, lecturers)
- Viewers (film enthusiasts, cinephiles)

It can be expected that professionals and viewers are the least familiar with data visualizations, graphs etc. but have substantial knowledge and interest in the content and effects of the videos. They would benefit the most from visualizations only showing data about a few movies with more detailed information about them. Such visualizations would allow them to gain insights about individual movies and analyze techniques used.

Academics are very familiar with plots and data visualizations and could be, on top of video-specific visualizations, also interested in aggregated data visualizations supporting the analysis of trends, patterns in the data set as a whole. In order to have adequate visualizations for academics, the visualizations should clearly indicate the meaning of numbers and scales as well as filters currently applied.

Task/Goals

- For film historians, it could be insightful to visualise the evolution of film techniques over time, showcasing which techniques emerged, evolved, or fell out of use across different periods. This mostly concerns image type (like computer or stop-motion animation or graphics) of sequences and to a lesser degree transitions (e.g. dissolve) or other attributes (e.g. rotoscoped). A research question in this context could be if there is an observable increase in (3D) computer animations in later years and more still images/live action in earlier years in the given data set.
- The application should help users understand how different film styles evolved over time, highlighting which types of films were most commonly produced in each style during different periods. The focus here is on the different genres of the videos created in the various time frames.
- Enable users to explore the various art styles of films produced across different countries, offering insights into regional cinematic trends and stylistic preferences. This task is aiming at identifying possible stylistic and contextual differences in animations from Austria, East Germany, and West Germany, e.g. are the technologies

used in different countries varied (with the GDR using less [computer] animation, 3D)?

- For film historians it could be interesting to identify which organisations were responsible for producing videos with certain technologies at a certain time and in which quantities to categorise the organisations by capabilities and influence.
- For film enthusiasts it could be interesting to identify the individuals involved in certain productions, e.g. director or animator to enable deeper exploration of their contributions and facilitate discovering their later works and broader career trajectories.

Visualization Prototype

In this section we describe how the visualisation prototype can be implemented addressing the different aspects of the Visualization Reference Model. In our approach we tried to follow the Visual Information Seeking Mantra: Overview first, zoom and filter, then details-on-demand.

Data Transformations

For reasons described in the data section production/release date requires cleanup. Furthermore, we reduce the dates to the respective year, because for the sake of the visualization smaller granularity is irrelevant.

To combine data from the two tables of videos and segments, we link and fuse the data using the `video_id` field.

To calculate segment duration we process the beginning and end timestamps.

In the segments table the same fields (Tier and Annotation_label) are used to describe Image Types, Transitions and Miscellaneous Attributes. We consider the Image Type most relevant for the users to fulfil the described tasks. Therefore, we identify Image Type from fields in the segments, while discarding Transitions and other attributes.

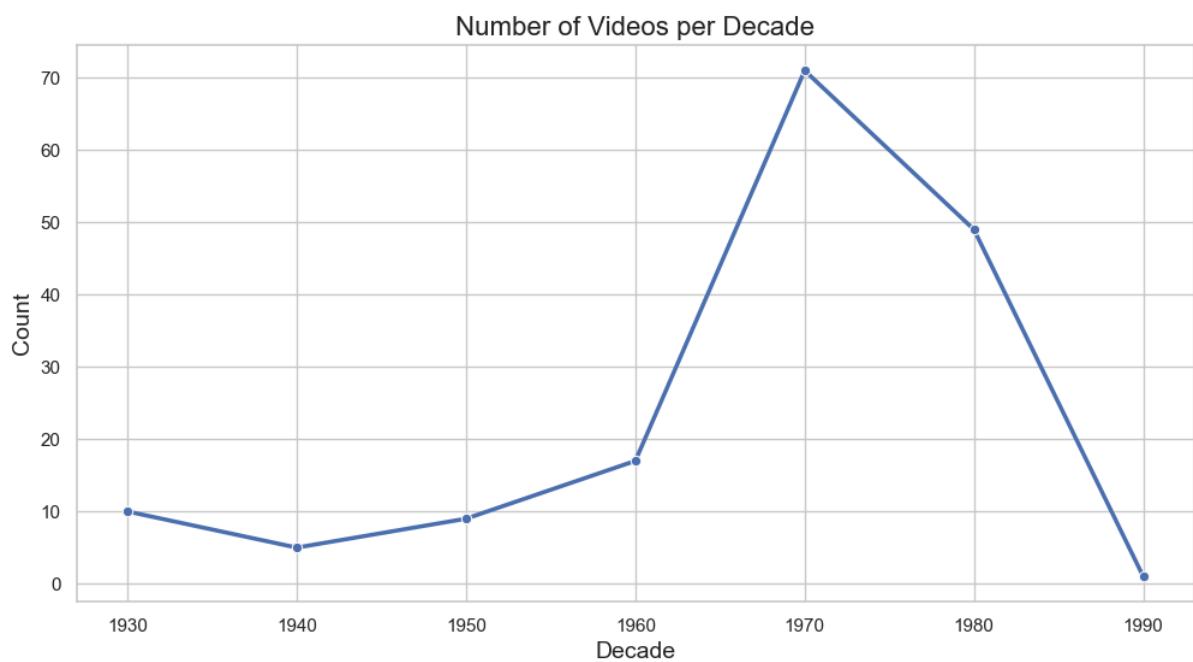
Visual Mappings

Line Graph

The line graph illustrates trends in video production and content characteristics over time. The position of each data point reflects the number of videos produced, organized by year or decade, or the duration of video segments for selected data. This mapping enables an analysis of temporal patterns and content distribution trends in the chosen dataset.

Mapping

Data Variable	Visual Variable
Decade (Quantitative)	X-axis position
Count of videos (Quantitative)	Y-axis position
Number of videos (Quantitative)	Line height and connected points

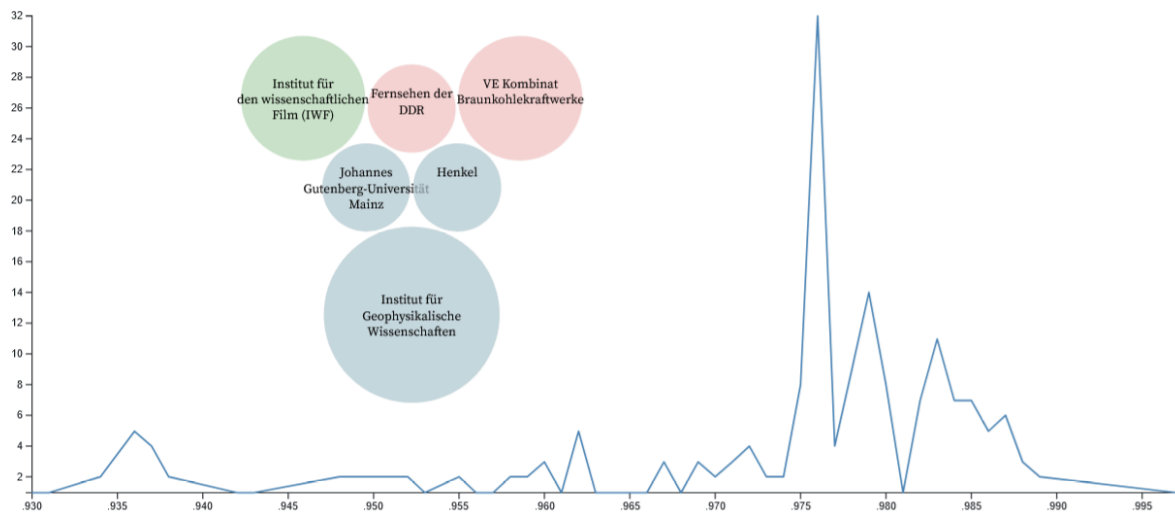


Bubble Chart

This bubble chart that is superimposed on the line graph visualizes production organizations/agents based on their geographic location and volume of film production. Each bubble represents an organization, with the bubble's size proportional to the number of films produced and its color corresponding to the organization's geographic region. This allows for an intuitive comparison of production output across different regions.

Mapping

Data Variable	Visual Variable
Time/Decade (Quantitative)	X-axis position
Number of occurrences (Quantitative)	Y-axis position (line chart)
Institutions/Entities (Nominal/Categorical)	Circle labels and grouping
Size of contribution/impact (Quantitative)	Circle size (area)
Category of institution/entity (Nominal/Categorical)	Circle color



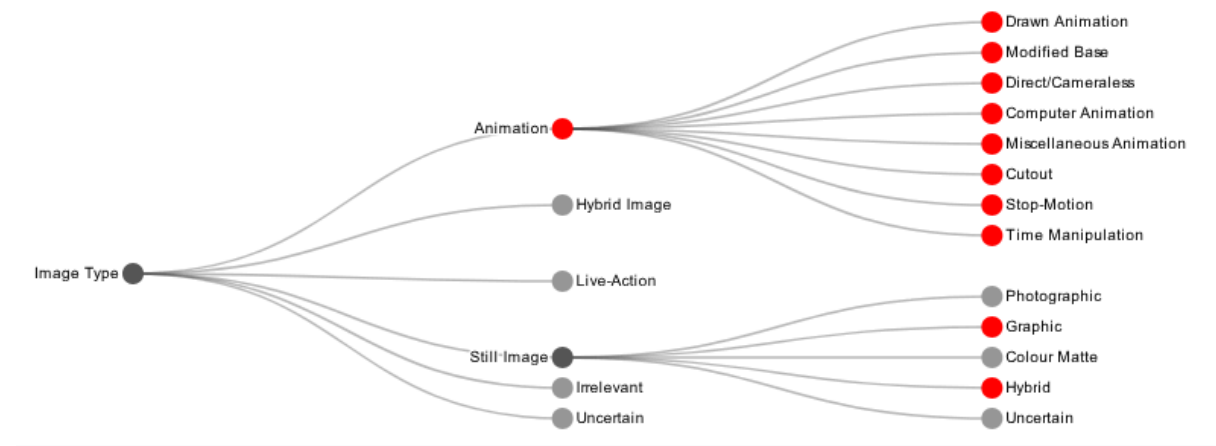
View Transformations

The main transformation possible for a user would be the possibility to zoom in and out to get more granular timeline views. In some cases aggregations over decades or multiple years might be useful. In other cases views of a few years could be more suitable.

Interactions

Filter

Selection of relevant image types can be done in a collapsible tree. In the current prototype clicking selects a node, while holding Shift while clicking expands a node. Selecting a parent node selects all children nodes as well. The advantage of this view is implicitly explaining the subcategories of image types. The image below illustrates the expanded tree with the selected nodes in red, while in the appendix you can also find images of the collapsed tree.



Other groups like country or genre could also be filtered using suitable UI elements like dropdowns or checkboxes.

Detail Views

Year Selection

By hovering or clicking on the graph detailed counts could be displayed for the specific years giving insights in the groups contributing to the total count. Interesting detail attributes could be organizations, contributing people, or genres.

Video Search

To show an individual movie or movies produced by a person in the time context, simple search fields could be used to select certain data points in the list of filtered videos. The specific elements (videos) could be plotted as points on the line graph.

Appendix: Preliminary Data Exploration

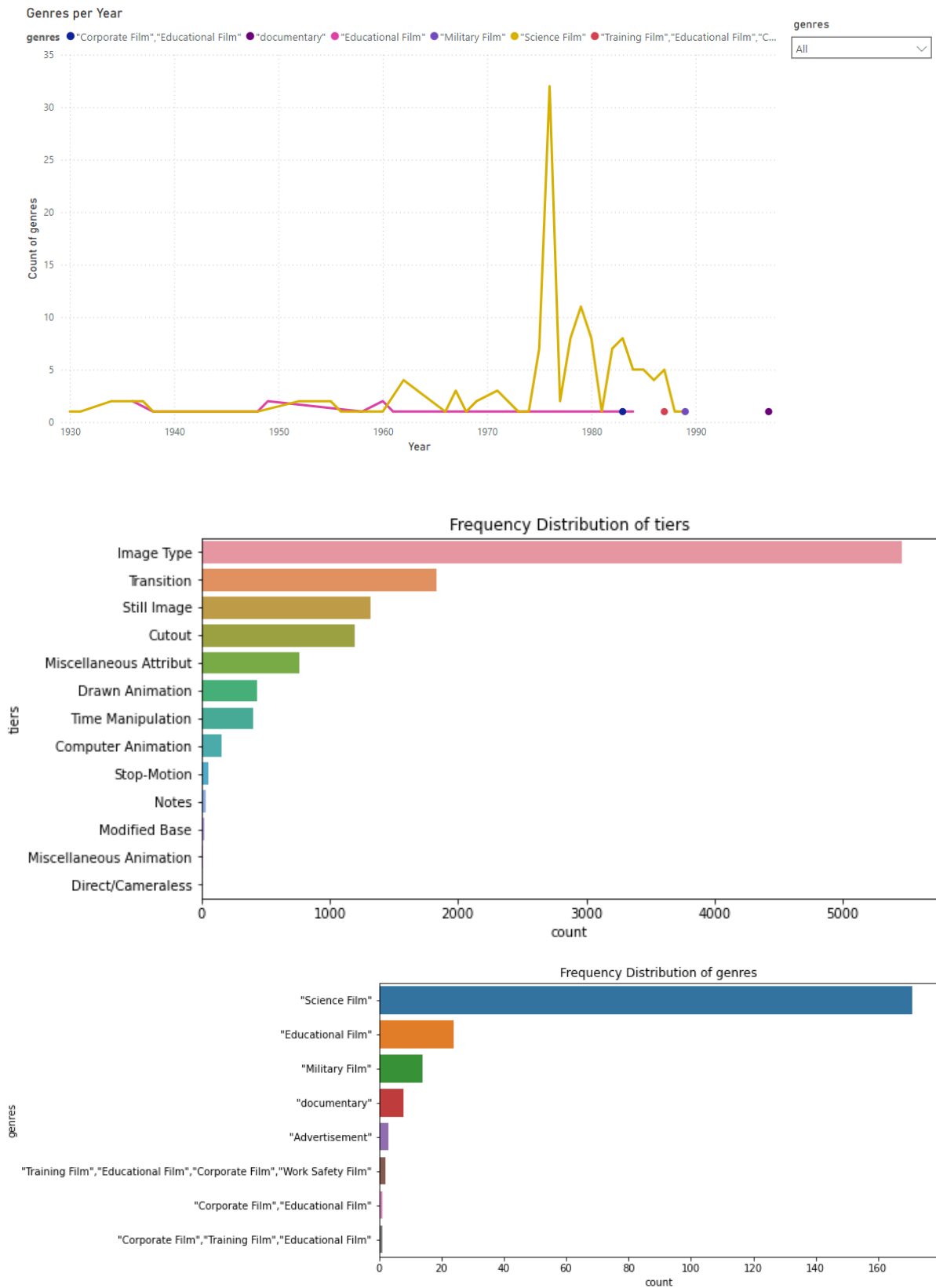


Image Type Filter Tree

Image type filter tree in two collapsed stages:

