Time-based Impact Mosaics

Stefan Jänicke*
Leipzig University, Leipzig, Germany

Figure 1: 390 Google Arts & Culture paintings with a landscape paper size format. In the 17th century the format was often used for biblical paintings and in the 19th century it was used for landscapes and genre paintings.

ABSTRACT

A Time-based Impact Mosaic is a photographic mosaic composed out of images of data items that have a certain impact on an observed topic. Images are scaled according to impact, so that important data items stick out. While the composition of all images resembles a quantitative time graph highlighting significant periods, individual data items that contribute to the development of a topic are explorable. Two usage scenarios on movies and paintings outline the effectiveness of Time-based Impact Mosaics.

1 INTRODUCTION

The proverb "A picture is worth a thousand words"—used for the first time in this phrasing exactly 100 years ago in a 1918 newspaper advertisement for the San Antonio Light's Pictorial Magazine of the War—is known as one main driving forces of visualization research. One of the characteristics of visualization is its interactivity necessary to gain diverse views on the data, to enable filtering or zooming operations, and to view details on specific data items [11]. As opposed to visualization, the power of infographics is that all information and knowledge that wants to be conveyed is quickly and clearly visible, without the need to interact [12]. However, visualizations often resemble infographics enhanced by means of interaction. For example, ThemeRiver [6] or Stacked Graphs [3] compare how quantities of different data categories changed over time like Rand McNally's Histomap (1931) that similarly contrasts the relative power of states, nations and empires in 4,000 years of history.

*e-mail: stjaenicke@vizcovery.de

This work is inspired by Adams' *Synchronological Chart* [1]. Adams juxtaposed biblical and world historical events from 4000BC to the 19th century, and he iconographically exposed important events or influential individuals. The size of an image represents the importance of an event, e.g., sigificant events such as the crucifixion of Christ reserve a large amount of space. Time-based Image Mosaics borrow this idea by scaling images according to their impact on the observed topic, and the result looks like a photographic mosaic [4]. Looking at the whole picture, the Time-based Image Mosaic communicates the temporal development of the observed topic, e.g., when impactful events took place or when periods with many significant events were. When taking a closer look, the bits and pieces that contribute to this development can be inspected.

This paper outlines two examples illustrating the capability of Time-based Image Mosaics to communicate developments of topics on the basis of image collections. First, movie posters are composed, so that conclusions on the development of the film industry or individual genres can be drawn. Second, a collection of paintings is analyzed according to typical canvas format sizes.

2 RELATED WORK

Temporal information is inherent in many data collections, so that numerous visualizations have been developed in order to support the analysis of trends. Different static time charting methods to illustrate quantitative and qualitative temporal information exist [5], and an overview of sophisticated interactive time-based visualization techniques is given by Aigner et al. [2]. One of those techniques are Stacked Graphs [3] that are used to illustrate topical changes in a dataset. The original idea introduced as the ThemeRiver by Havre et al. [6] aimed at visualizing thematic variations over time within large document collections. Shi et al. [10] enhance such graphs by placing predominant tags at certain positions in streams. Time-based Image Mosaics are similar as images are scaled according to impact.

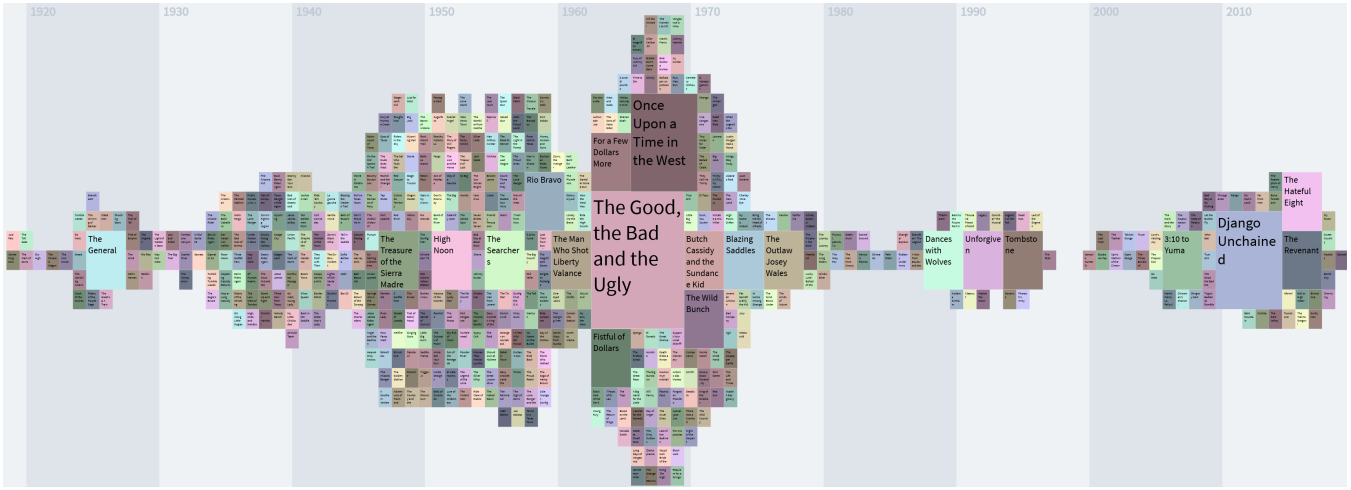


Figure 2: Time-based Image Mosaic of IMDb western movies. Different periods of the genre can be identified: (1) many classical western movies were produced in the 1950s, (2) the spaghetti western era in the 1960s with some movies having a high impact but a smaller number of movies with lower impact values, (3) the vanishing genre in the 1970s, and (4) the two neo western periods in the 1990s and 2010s. *Due to copyright issues, poster thumbnails needed to be replaced by colored rectangles.*

Three related works also place images on a horizontal temporal axis. Huynh et al. [7] arrange images on a timeline in columns, and representative images are scaled to facilitate semantic zooming. The size of an image does not represent its individual impact but the importance of the cluster it belongs to. In addition, the arrangement of images leads to whitespaces, and the timeline does not reflect the importance certain time ranges. Itoh et al. [8] stack image thumbnails extracted from blogs and TV on a 3D timeline. While time-dependent quantities are conveyed, the importance of an image is not outlined as all images have the same size. Timages [9] is similar to Time-based Image Mosaics. It takes images with arbitrary aspect ratios, and it provides a space filling approach to place thumbnails on a timeline. This work is complementary to Timages as it minimizes white spaces when all images have the same aspect ratio.

3 Design

A Time-based Image Mosaic arranges images of data items on a horizontal timeline. In the following, let $D = \{d_1, \ldots, d_n\}$ denote a data set with n data items. The following constraints are given:

- The image of a data item d_i needs to intersect the vertical axis of the corresponding time stamp.
- To visualize data items having a high impact to the observed topic more prominently, images are scaled dependent on impact values. The image of d_i is scaled according to d_i's impact I_i—a positive floating-point number in the range $(0,1]$.
- As all images have the same size, thus, a regular grid is used to place the images aiming to minimize white spaces.
- The images of the data items having the highest impact shall be placed first, thus, D is sorted by decreasing impact values. So, d_1 stands for the data item with the highest impact, and d_n stands for the data item with the lowest impact.

Image Scaling: D is divided into a user-defined number of k impact classes $c_1, \ldots, c_k$. The class $c(d_i)$ of data item d_i is defined by

$$c(d_i) = \frac{I_i - I_n}{I_1 - I_n} \cdot (k-1) + 1.$$

Sizes of images depend on the granularity of the underlying timeline. The width of an image belonging to the least important class c_1 is set to the width w_{min} of the shortest time unit of the timeline. An image of another class c_j gets width $(2j-1) \cdot w_{min}$. Thus, if the smallest time unit is a year, an image of class c_1 will cover exactly one year, an image of class c_2 will cover three years, etc.

Image Placing: Images are placed as close as possible to the horizontal center of the timeline, and an image of data item d_i is required to intersect the vertical line $x(t_i)$ of the timeline that corresponds to the time unit d_i belongs to. Images are moved until no overlaps with already placed images occur. After placing $i-1$ images of D_{i-1}, the first attempt to place the image of d_i is at position $\{x(t_i), 0\}$. If it overlaps any image of D_{i-1}, horizontal movements alternating to the left and to the right by multiples of w_{min} are tested as long as d_i intersects the vertical line $x(t_i)$. If no valid position is found, the y-position then alternates upwards and downwards by multiples of h_{min} that is the height of images of class c_1. For each new y-value, all possible x-values are tested as described above. The algorithm terminates if valid positions are found for all $d_i \in D$.

4 Use Cases

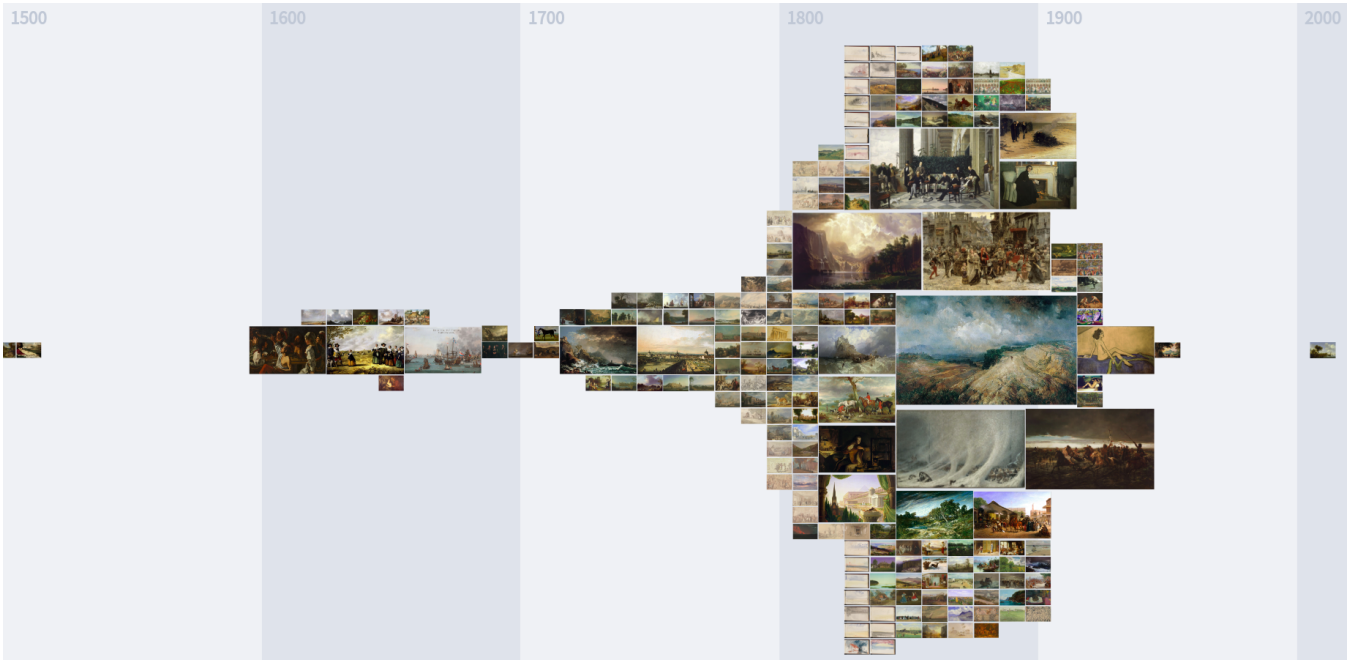
Two scenarios illustrate the capability of Time-based Image Mosaics to generate visually appealing, infographic-like time-based visualizations. In all examples, a number of 5 classes was used. Dependent on the scenario, "impact" on the observed topic is individually defined.

Movie Posters. The Internet Movie Database (IMDb)[1] contains information about more than 180,000 movies. For arranging the posters of movies having high impacts on a timeline, the impact I_i of a movie d_i is defined according to IMDb's weighted rating (also used to define the IMDb Top250) as

$$I_i = \frac{v_i}{v_i + 25{,}000} \cdot r_i + \frac{25{,}000}{25{,}000 + v_i} \cdot 7$$

with the number of votes v_i for the movie, and the average rating r_i. Movies with at least 5,000 user votes are considered. Figure 2 provides an analysis of the western genre, and Figure 4 illustrates the changing impact of the horror genre.

[1] http://www.imdb.com/



(a) 211 paintings with a horizontal golden ratio format that was often used in the 19th century. Typical subjects of the paintings are sea- and landscapes.

(b) 304 paintings with a vertical golden ratio format. Nearly all paintings are portraits.

Figure 3: Google Arts & Culture paintings with a golden ratio format.

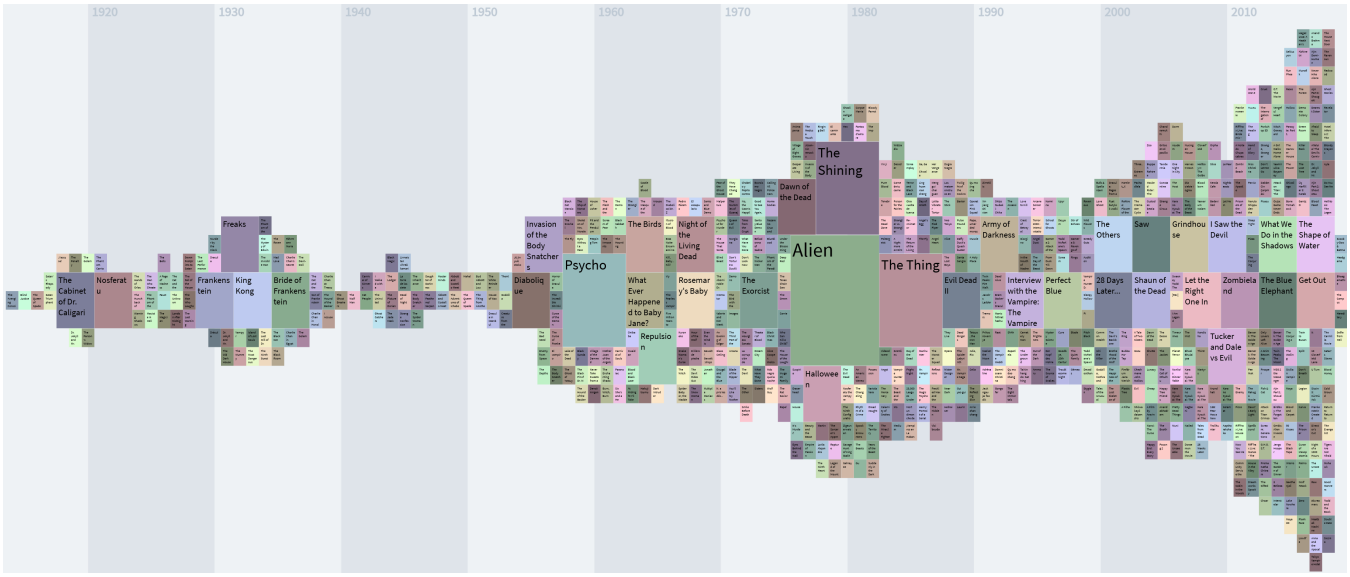


Figure 4: Time-based Image Mosaic of IMDb horror movies. The heyday around 1980 with important movies like "Alien (1979)" and "The Shining (1980)" stick out. Furthermore, milestones of the genre like "Nosferatu (1922)" and "Psycho (1960)" are prominent. *Due to copyright issues, poster thumbnails needed to be replaced by colored rectangles.*

Paintings. Google Arts & Culture[2] is an online platform that provides access to high-resolution images of artworks. Around 9,000 images of paintings linked through Wikipedia were crawled, and the contents of paintings in standardized canvas formats were analyzed. For illustration purposes, sizes of paintings are mapped to impact in these examples. Thus, the larger a painting the larger its thumbnail in the mosaic. A typical format complied to the golden ratio with a width/height ratio of approximately 1.618. Paintings having that format are shown in Figure 3a. The format was often used in the 19th century, and typical subjects of the paintings are sea- and landscapes. Paintings having a vertical golden ratio format almost exclusively show portrait paintings (see Figure 3b). Paintings having today's standard paper size format with a width/height ratio of approximately 1.414 are shown in Figure 1. In the 17th century the format was often used for biblical paintings, and in the 19th century it was used for landscapes and genre paintings.

5 Summary

The prior goal of Time-based Image Mosaics is giving a visually appealing, infographic-like overview of the development of a topic of a time-stamped data collection in the form of a time graph composed of corresponding images. However, Time-based Image Mosaics can be used in an interactive Web-based environment as zooming and filtering operations are possible, and details about an individual data item are shown on demand. Time-based Image Mosaics are implemented as a user-configurable JavaScript library based on D3.js.[3]

[2]https://artsandculture.google.com/

[3]http://timages.vizcovery.de

References

[1] S. C. Adams. Chronological Chart of Ancient, Modern and Biblical History, 1871. `https://en.wikipedia.org/wiki/Adams_Synchronological_Chart_or_Map_of_History` (Retrieved 2018-02-28).

[2] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of Time-Oriented Data*. Springer Publishing Company, Incorporated, 1st ed., 2011.

[3] L. Byron and M. Wattenberg. Stacked Graphs – Geometry & Aesthetics. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1245–1252, Nov 2008.

[4] A. Cartwright. *Mixed Emulsions: Altered Art Techniques for Photographic Imagery*. Quarry Books, 2007.

[5] R. L. Harris. *Information Graphics: A Comprehensive Illustrated Reference*. Oxford University Press, Inc., New York, NY, USA, 1999.

[6] S. Havre, E. Hetzler, P. Whitney, and L. Nowell. ThemeRiver: Visualizing Thematic Changes in Large Document Collections. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):9–20, Jan. 2002.

[7] D. Huynh, S. M. Drucker, P. Baudisch, and C. Wong. Time Quilt: Scaling up Zoomable Photo Browsers for Large, Unstructured Photo Collections. In *Conference on Human Factors in Computing Systems (CHI'05)*, pp. 1937–1940. ACM New York, NY, USA, 2005.

[8] M. Itoh, M. Toyoda, C. Z. Zhu, S. Satoh, and M. Kitsuregawa. Image Flows Visualization for Inter-media Comparison. In *2014 IEEE Pacific Visualization Symposium*, pp. 129–136, March 2014.

[9] S. Jänicke. Timages: Enhancing Time Graphs with Iconographic Information. In *Leipzig Symposium on Visualization in Applications (LEVIA'18), Leipzig, Germany*, October 18th, 2018.

[10] L. Shi, F. Wei, S. Liu, L. Tan, X. Lian, and M. Zhou. Understanding text corpora with multiple facets. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pp. 99–106, Oct 2010.

[11] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages*, pp. 336 –343, 1996.

[12] M. Smiciklas. *The Power of Infographics: Using Pictures To Communicate and Connect With Your Audiences*. Que Publishing, 2012.