MoVis - Movie Recommendation and Visualization

CPSC 547 Infomation Visualization Project

Ye Chen - clara.yechen@gmail.com

Yujie Yang - yangyujie.hust@gmail.com

Previous work

Visualization Tools for Media Libraries

Several visualization tools have been designed for browsing, exploration, and searching through large collections of movies, music, photos, books[1, 2, 4, 3, 5, 6]. Among those visualization tools, some of them implement a scatterplot-based main view[1, 2, 3, 6]. FilmFinder [1] features tightly coupled interactive filtering, where the result of moving sliders and pressing buttons is immediately reflected in the visual encoding. Musicovery [2] is an online music streaming tool that visualizes its music collection based on mood and emotional content of music tracks. The interface makes good use of space and color coding in various filter types including genre, date, artist, intended activity,and etc. PhotoFinder [3] is designed to be easy to use for searching and browsing photos by using a scatter plot thumbnail display and drag-and-drop interface. Bohemian bookshelf [6] is a visualization interface for browsing a book collection with the goal of supporting serendipity. The interface includes five interconnected views that encode physical and/or semantic properties of the books in various ways. The interface is designed to support playful interaction and provide various access points to explore the collection.

However, other tools proposed some other features: treemap, disc or rectangle visualization as their main view[4, 5]. Torrens [4] proposed three visualizations for personal music collections: disc, rectangle and treemap and he suggested to use treemap visualization. Pocket PhotoMesa [5] is an image browser for the pocket pc that employs quantum strip treemaps for laying out images and zoomable user interfaces for navigation.

Apart from the usual visualization approach, some media visualization systems use machine learning algorithms such as clustering to group and present media items based on their similarity [7, 8]. Because of limit of time and goal of the MoVis, we don't focus on using machine learning algorithms on this stage.

Our main page uses interactive scatterplot for exploration and shares characteristics with some of the above examples. Those examples provide us with support of using scatterplot-based main view for visualization of media libraries.

Social network Visualization

Visualization of social networks has a rich history, particularly within the social sciences, where node-link depictions of social relations have been employed as an analytical tool since at least the 1930s. Linton Freeman documents the history of social network visualization within sociological research, providing examples of the ways in which spatial position, color, size, and shape can all be used to encode information [9].

In recent years, such approaches have been effectively used in the analysis of domains such as e-mail communication [12], early online social networks [13], and co-authorship networks in scientific publications [14]. There are a number of systems for generating such social network visualizations and performing statistical analyses for the purpose of sociological research, such as JUNG [15].

In addition to sociological research, there have been numerous social visualization projects intended for end-users rather than outside researchers. Vizster[10] is a visualization tool for playful end-user exploration and navigation of large-scale online social networks. It builds upon familiar node-link network layouts to contribute customized techniques for exploring connectivity in large graph structures, supporting visual search and analysis, and automatically identifying and visualizing community structures. Co-author graph in Microsoft academic search [11] shows us the top 30 co-authors of the main researcher in the middle. It allows users to drag the nodes to explore the connection and change the main researcher in the middle. TouchGraph [16] uses a force-directed layout to present a network visualization of users of the LiveJournal online community, allowing personal networks to be expanded or contracted by user interaction.

The actor/director connection graph in MoVis shares lots of features with the above visualization tools, especially Vizster [10] and the co-author graph [11].

Update(Nov. 26)

1. A new data set has been scrawled from http://www.imdb.com/search/title?at=0&sort=boxoffice_gross_us&start=1&title_type=feature&year=1995,2014, with 104181 data in total. Compared to the original dataset we have, this new dataset has more attributes (like lists of actors for each movie, language and countries) and larger data size. We might use some of the new attributes in MoVis later on.
2. Cleaned the dataset in Google Refine:
    1. Reduced size of "genre" types from 23 into 14 categories by grouping same genre together (For example, grouping "Drama" "Drama " "Dram" into "Drama").
    2. Formatted "budget" into US dollars for all movies, given the budget for some movies are in other currencies. The number of currency types were more than 25, including HUF, AUD, NOK, HKD, HRK, RUR, MYR, MXN, BRL, TRL, ILS and so on.
    3. Formatted releaseDate into "%d %m %y".

4. Split the actors & genre & country where more than one data item is shown in one cell into multiple cells in the same row.

3. Built basic scatterplot of movie data with rating as Y-axis and releaseDate as X-axis as shown in Figure 1, where tooltip about basic information of movie will be shown when hovering over a data point. (This figure is still using our original data, so it only contains around 1000 data points. We will change the dataset afterwards)

4. Coloured each data point according to its genre.

5. Divided boxoffice into three classes and encoded them into the size of the points.

6. Added zoom-in and zooming function and dragging function into the scatterplot, as shown in Figure 2.

7. Added in the pop-up function (Figure 3) where a new window pops up with lists of movie information when a point is clicked on the scatterplot. In the new window, when clicking "link to iMDB", it goes directly to iMDB page of the chosen movie(Figure 4).

8. We also played with filter function to show only the movies with boxoffice greater than 5000.

9. We tried to rescale the X and Y axis while keeping the location unchanged before zooming in. We tried to follow the example in http://bl.ocks.org/stepheneb/1182434, but we haven't found a good solution yet.
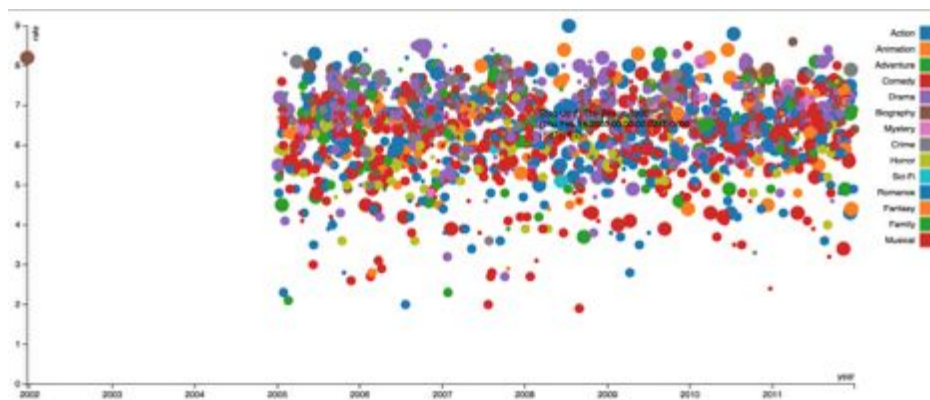

Figure 1 Scatterplot view in main page
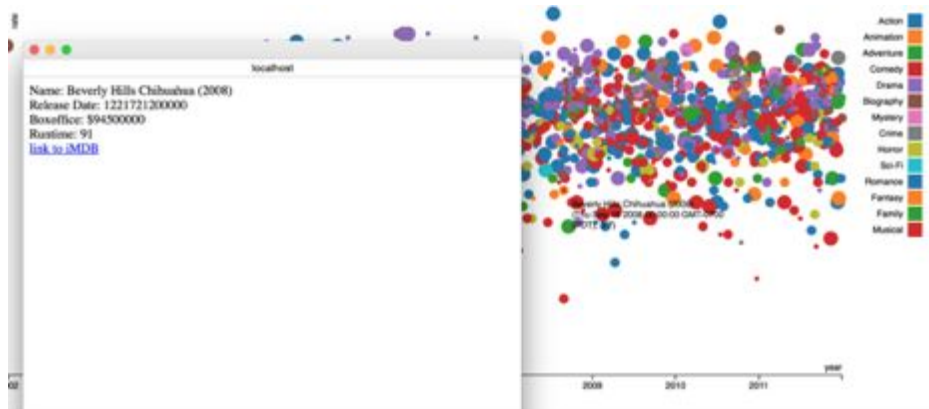

Figure 2 Zooming function in the scatterplot

Figure 3 Pop-up window in the scatterplot

Figure 4 When clicking Link to iMDB, the corresponding iMDB page shows up

Split of the work

Previous work: Yujie Yang
Update: Ye Chen

Reference

[1] Ahlberg C, Shneiderman B. Visual information seeking using the filmfinder[C]//Conference companion on Human factors in computing systems. ACM, 1994: 433-434.
[2] Musicovery. [http://musicovery.com]
[3] Kang H, Shneiderman B. Visualization methods for personal photo collections: Browsing and searching in the photofinder[C]//Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on. IEEE, 2000, 3: 539-1542.
[4] Torrens M, Hertzog P, Arcos J L. Visualizing and Exploring Personal Music Libraries[C]//ISMIR. 2004.
[5] Khella A, Bederson B B. Pocket PhotoMesa: a Zoomable image browser for PDAs[C]//Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia. ACM, 2004: 19-24.
[6] Thudt A, Hinrichs U, Carpendale S. The bohemian bookshelf: supporting serendipitous book discoveries through information visualization[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2012: 1461-1470.

[7] Moghaddam B, Tian Q, Lesh N, et al. Visualization and user-modeling for browsing personal photo libraries[J]. International Journal of Computer Vision, 2004, 56(1-2): 109-130.

[8] Nguyen G P, Worring M. Interactive access to large image collections using similarity-based visualization[J]. Journal of Visual Languages & Computing, 2008, 19(2): 203-224.

[9] Freeman L C. Visualizing social networks[J]. Journal of social structure, 2000, 1(1): 4.

[10] Heer J, Boyd D. Vizster: Visualizing online social networks[C]//Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on. IEEE, 2005: 32-39.

[11] Co-author graph [http://academic.research.microsoft.com/VisualExplorer]

[12] Fisher D, Dourish P. Social and temporal structures in everyday collaboration[C]//Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 2004: 551-558.

[13] Adamic L, Buyukkokten O, Adar E. A social network caught in the web[J]. First monday, 2003, 8(6).

[14] Newman M E J. Coauthorship networks and patterns of scientific collaboration[J]. Proceedings of the national academy of sciences, 2004, 101(suppl 1): 5200-5205.

[15] JUNG: Java Universal Network/Graph Framework. [http://jung.sf.net/]

[16] Touchgraph [http://touchgraph.com]