A Quantitative Study of the Geographical Factors in the Mainland-Chinese Film Industry, 1949 -1966

Video Document

1. Filmography Visualisation (00:00-00:35)
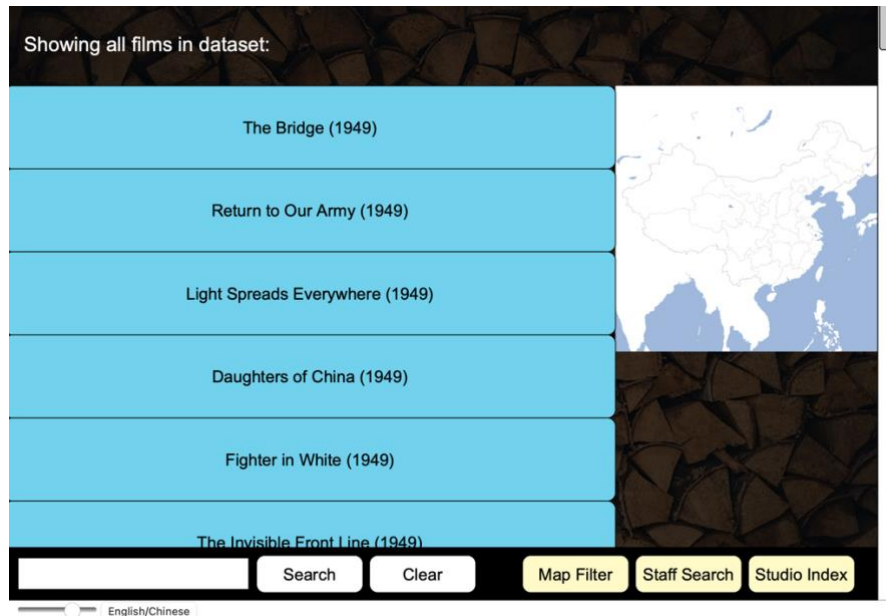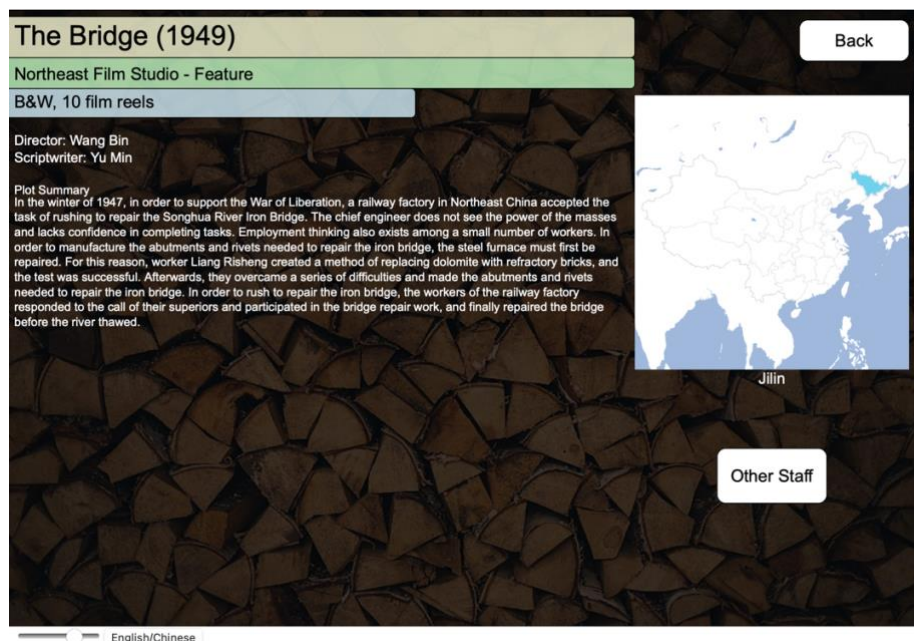

Figure 1.1. Filmography menu.

Figure 1.2. Film information.

This part of the video demonstrates the user interface of the filmography visualising program used in my project (*Figure 1.1*).[1] This program is used to improve the accessibility of the metadata I collected for the project. In this menu, users can search for films in Chinese or translated English titles. Clicking the button of a film will direct the user to this film's information panel (*Figure 1.2*). This panel documents the metadata of the film and shows the location of film production on the map on the right side of the interface.
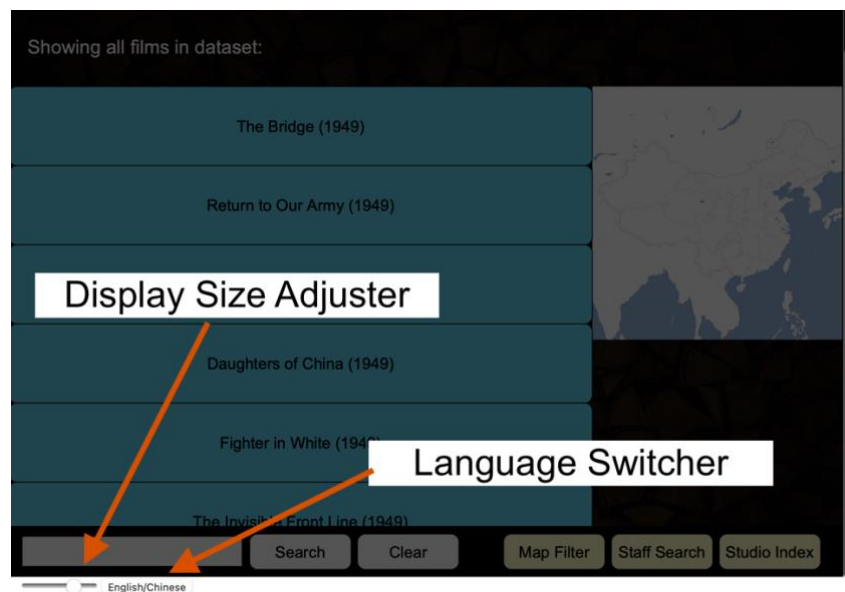

Figure 1.3. Two display setting components.

On the left bottom corner of the interface, a display size adjuster and a language switcher are added (*Figure 1.3*). The display size adjuster is used to set the interface size relative to the window, and the language switcher allows the user interface to switch between English and simplified Chinese in real time.

---

[1] The program can be accessed at: <https://filmpreservation.github.io/ChineseCinema/Visualiser/main.html>.

2. Map Searching (00:35-01:03)
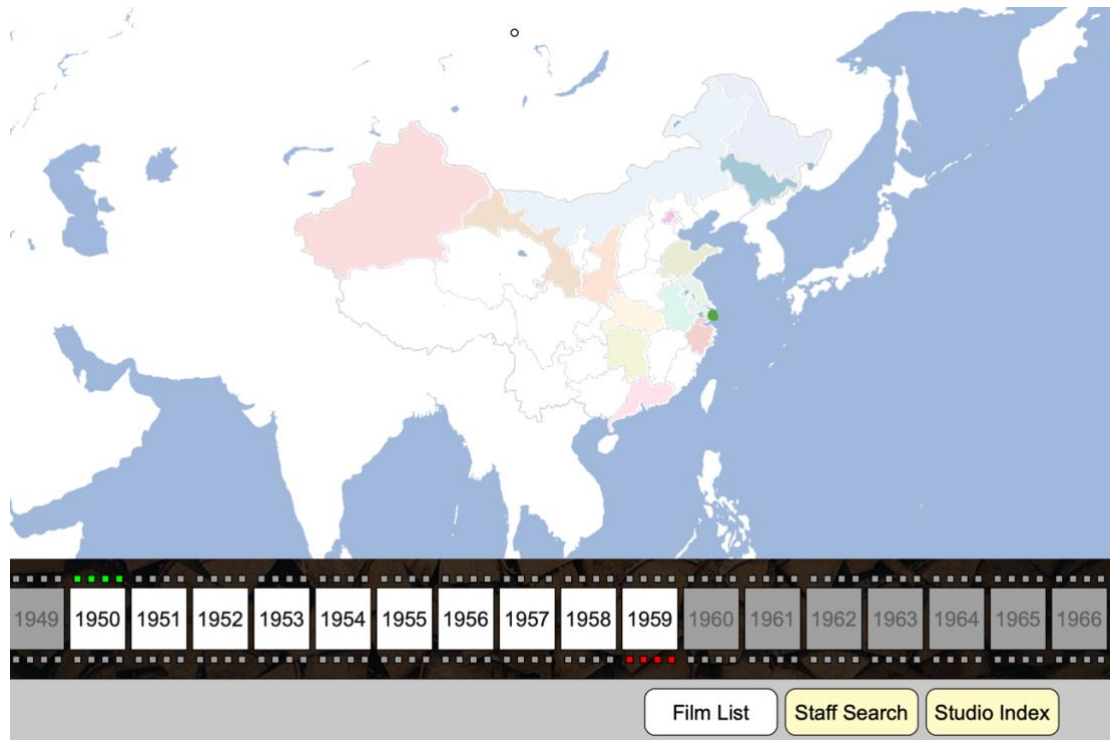
Figure 2.1. The map searching interface.

Map searching is an accessible feature included in the filmography program (*Figure 2.1*). It allows the user to search all films produced in a region within a selected range of years. The user can click on the film reel at the bottom of the map to set the year range and then select a region on the map to retrieve the films produced in the selected region and years.
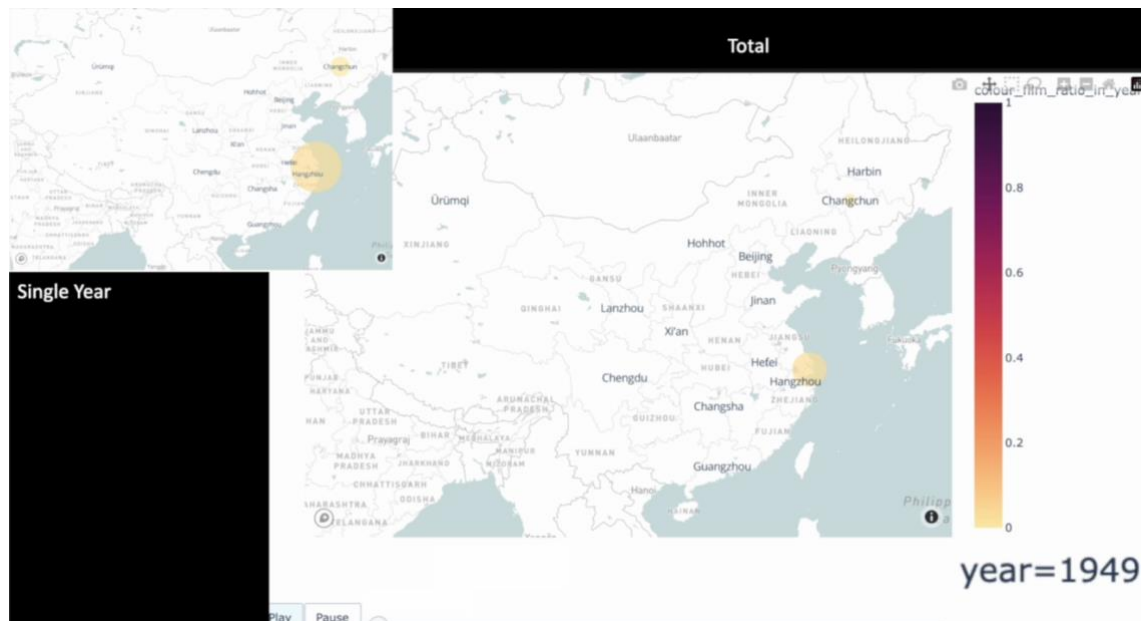
3. Filmmaking Regions (01:03-01:30)

Figure 3.1. Visualisation of filmmaking regions.

This part of the video shows the visualisation of the filmmaking regions based on the Geographical Information System library Mapbox (*Figure 3.1*). In this visualisation, the size of circles is used to represent the number of films produced in a region, and gradient colours represent the proportion of colour films produced in the region. The larger graph illustrates the total number of films produced in a region from 1949 to the displayed year, while the smaller one in the top left corner shows the specific number of films produced in the year.

4. Film Practitioner Movements (01:30-02:00)

Figure 4.1. Visualisation of filmmaking regions.

This section of the video visualises the "GeoCategory" algorithm used in this project (*Figure 4.1*). Based on the locations of the film studios with which the film practitioners were affiliated, these film practitioners are mapped to geographical areas. It can illustrate personnel changes in the film industry and the expansion of state-owned studios. In order to observe the movements of former members of Shanghai's private film companies, I marked people who worked with these companies in purple and other people from state-owned studios or non-Chinese institutions in green to differentiate them.

Figure 4.2. Paris filmmakers involved in the network of the Chinese film industry.

This visualisation also shows some practices of transnational film production. For example, when Beijing Film Studio and Garance Productions in Paris co-produced the film *Cerf-volant du bout du monde*, some French filmmakers can also be seen in this visualisation program (*Figure 4.2*).

5. Topic Modelling (02:00-02:30)

Figure 5.1. Topics modelling visualisation.

This part of the video shows a topic modelling visualisation generated with Gensim and pyLDAvis. Gensim extracted sixteen topics from the 732 Chinese films (*Figure 5.1*).[2] These topics also show the importance of ideology education in Chinese cinema between 1949 and 1966.

---

[2] For the English translation of the results, please see:

<https://github.com/FilmPreservation/ChineseCinema/blob/main/Topic/Topics_Summarised.csv>.

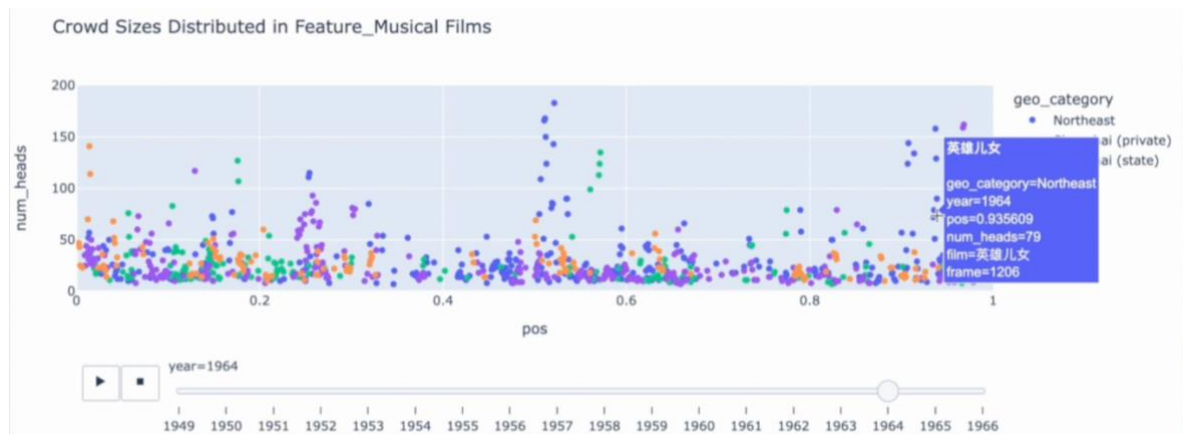## 6. Evolution of Crowd Size and Smile Intensity (02:30-03:35)

*Figure 6.1.* Crowd size distribution in 1964.

Using the scatterplot functionality provided by the Plotly library, I designed this visualisation program to combine quantitative and qualitative analyses (*Figure 6.1*). It takes the size of a crowd or the intensity of a face's smiling identified by computer vision techniques as the y-axis and the crowd or the face's position in the film's narrative time as the x-axis. This scatterplot allowed me to locate some typical crowd scenes and intense smiles effectively.

7. Face Clustering (03:35-04:42)

Figure 7.1. Face clustering graph.

The last section of the video presents my methodological practice, a human identification program based on face clustering. It uses the high-dimensional data generated by the Facenet512 model to represent facial features and uses the t-SNE algorithm to reduce the dimensionality and generate a two-dimensional graph (*Figure 7.1*). The graph is also interactive and editable. The user can remove unwanted faces from a cluster by using the editing mode implemented in the graph program.[3]
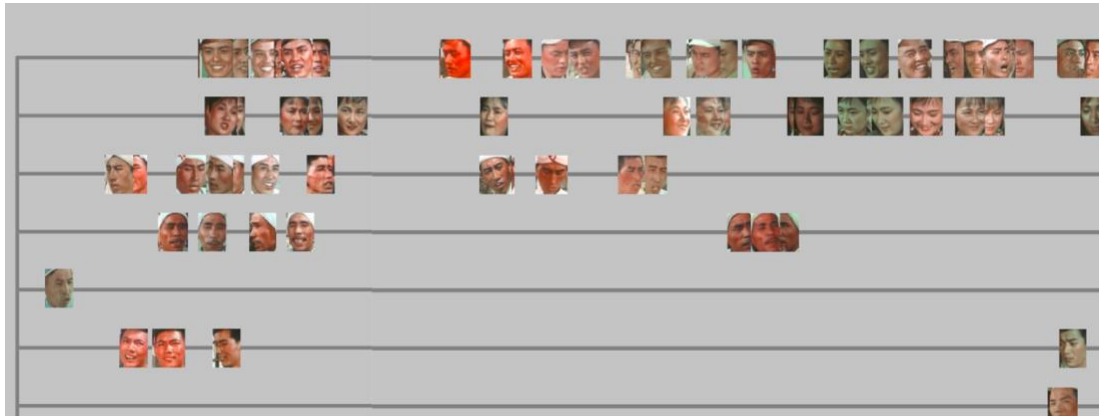
Figure 7.2. Face clustering in film timeline.

This graph program also implements a timeline function that allows the user to map the appearances of an identified entity to the timeline of the film. This feature might be useful for future research on narrative structure and character relations.

---

[3] For a detailed illustration of this program, please see: <https://github.com/FilmPreservation/FaceClustering>.