World Flags Similarity

Richard Liu*
Massachusetts Institute of Technology

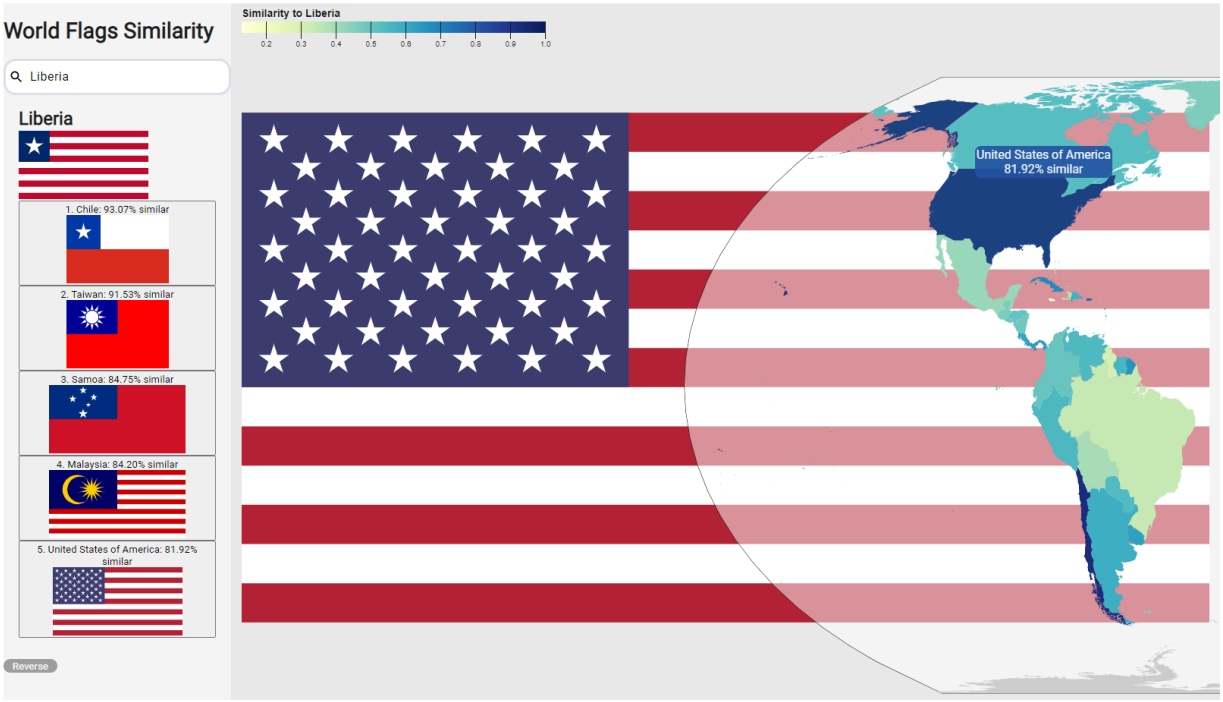

Figure 1: Demo of the World Flags Similarity application.

Abstract

Flag designs from around the world may appear simple, yet they can capture the rich history and culture of a country. This project aims to help the user learn something new about sovereign state flags, through the lens of flag similarity. By highlighting trends across flags and the history behind certain flag features, this project seeks to make flags more accessible and allow users to explore the data to gain deeper insights. The final interactive visualization, which offers a variety of interaction techniques, can be overwhelming for a new user. To this end, we also present a walkthrough tool which, inspired by scrollytelling techniques, introduces users to all of the features of the visualization.

Index Terms: H.5.m [Information Interfaces and Presentation]: Miscellaneous

1 Introduction

Flag designs from around the world may appear simple, yet they can capture the rich history and culture of a country. There are elements which are common to many flags, but also many elements which distinguish flags, and exemplify the unique history and cultures of the different sovereign states. This project aims, broadly, to enable users to discover these trends, through the lens of flag similarity. To this end, we present our work with primary focus on the problems of calculating flag similarity, and onboarding new users to the different features of an interactive visualization.

The problem of calculating flag similarity is to evaluate the relationship between the feature vectors representing two country's flags, and assign a non-negative score correlating with similar flag design. In particular, in developing this application we believed visualization of flag similarity to indeed be an effective means of learning about the elements of national flags and the relevant history and culture. Discussed more in the section on insights, we find that some countries with ties in their history and culture, indeed share elements of their flag design. And some shared elements of flag design, like the use of stars to represent a nation's administrative subdivisions, can inform a general sense of a country's history.

From user feedback on an initial iteration of the World Flags Similarity application, it was clear that the original prototype visualization was not always successful in "onboarding" new users to the different features offered by the application. To this end, and inspired by scrollytelling techniques and video game tutorials, we also propose and implement here a (perhaps novel) walkthrough system. In our design of this system, our aim was to enable the user's onboarding as well as build a system which may be reused to onboard users to any interactive system.

2 Background and Related Work

This section introduces some key datasets used for this project, and other research or findings which tie into this work.

*e-mail: rtliu@mit.edu

2.1 Datasets

The main dataset used is the Flags Data Set from the UC Irvine Machine Learning Repository [5], which represents the flags of various nations as a kind of feature vector (the dataset is current with the year 1986).

Images of flags and insights into flag design and history are sourced from Wikipedia [2, 3].

2.2 Other Related Work

Complementing the application developed in this paper, there exist several resources which document connections between national flags and the associated culture and history: [1, 6, 8]

A visualization with many interactions can be overwhelming for new users. There are several options to introduce a new users to all the features. Perhaps most simple is to use a simple, purely text-based approach akin to an instruction manual. Another approach, apparently popular in online journalism, is *scrollytelling*, which can be thought of as a semi-interactive instruction manual [10]. The approach we have implemented is a fully interactive walkthrough or tutorial system. There exists in the literature some evaluations of the interactive tutorial system [7, 9]. However, we believe that our approach is novel within the ecosystem of interactive data visualizations, and we will discuss the possible advantages and particular implementation of our walkthrough system.

3 Methods

3.1 Data Parsing

In order to source the correct images for each of the sovereign states of interest, I first use the Pandoc library [4] to parse the timeline of national flags, converting the HTML to a markdown format. From here, a Python script using string operations generates a resulting, clean dataset mapping each sovereign state to a vector flag image.

3.2 Flag Similarity

To calculate flag similarity, we represent each country's flag as a feature vector in a high dimensional space. Flag similarity is assigned as a non-negative score comparing two flags, and is the weighted sum of three different classes of features. For nominal features such as color, a score component of 1 is given for a match, and 0 otherwise. For boolean features such as "presence of text", letting a, b be the value of the two input features, we give a score component of $f_B(a,b) = 0.8 \cdot a \cdot b + 0.2 \cdot (a \Leftrightarrow b)$. For interval features like "number of stars", we give a score component of $f_I(a,b) = \frac{\min(a,b)}{\max(a,b)}$.

The choice of f_B resembles the nominal features score function, but relatively disfavors matches with a "False" value, only assigning a score component of 0.2 in this case. This seems natural, as two flags that might be "similar" in the manner of *not* having text present, should not necessarily be rewarded a full score component. The choice of f_I also resembles the nominal features score function, and has the seemingly natural property of scaling with the numerical value of the inputs. Overall, we found these choices to be effective in calculating flag similarity.

Since the flag similarity data is aggregated into geographic units of countries, the Chloropleth appears to be an effective encoding of the data of interest (see Fig. 2).

3.3 User Interactions

In the left panel (see Fig. 3), users can query for particular countries as well as view the top countries with similar flags to the queried country's flags. Additionally, users can opt to reverse the order of the list of similar flags, to instead view the least similar flags and potentially surface some new trends. Finally, users can click on a country in the list to query for it. This is a powerful interaction allowing the user to explore across chains of similar national flags.

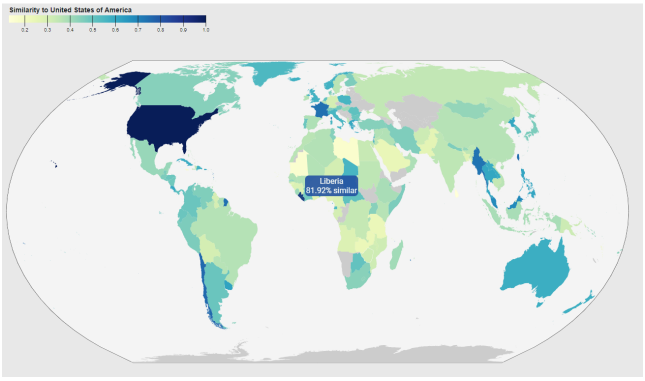

Figure 2: A visualization of the flag similarity data with respect to the U.S. flag.

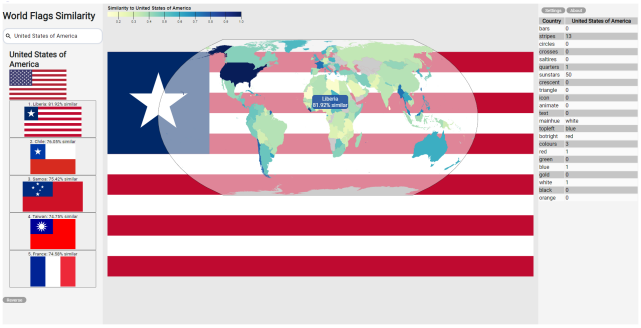

Figure 3: A visualization of the flag similarity data with respect to the U.S. flag.

In the middle panel, users can visualize flag similarity around the world, hover on a country to view the percentage of similarity and visually preview the flag, and pan and zoom on the map. Additionally, users can click on a country on the map to query for it. This is a powerful interaction allowing the user to explore flag similarity across geographic proximal nations.

In the right panel, users can visualize the feature vectors which describe a country's flag. Additionally, users can click on a feature in the table, which queries for it; that is, a detail query is performed which visualizes the trends of that feature in the flags across the world. For example, clicking on "stars" will show how many stars each country's flag has. This is a powerful interaction allowing the user to explore particular flag features across the world, and also activate *Insights* to learn about relevant history and culture. Finally, the user can adjust some settings and read information in an "About" section.

3.4 Walkthrough/Tutorial

We implement a walkthrough or tutorial system which prompts the user to take certain actions, and offers visual feedback as to the actions that are taken. We find the system to be modular in design, and potentially reusable across different interactive web applications. The system, for each step of the walkthrough, takes some inputs `walkthroughStep, requirementsSatisfied` and will increment the current `walkthroughStep` when the specified requirements are satisifed (e.g. after some element is clicked). In another module, the system can be specified to display particular prompts and feedback depending on the current `walkthroughStep`. It is our hope that, if the walkthrough system is found to be widely effective, may indeed be adapted and reused across different interactive visualizations and other applications.

In particular, the walkthrough in the World Flags Similarity application takes the user through the following:

1. In what country were you born? — User learns how to query by country name.

2. Which flag is most similar . . . ? — User learns interactions with the top similar countries panel.

3. Now, navigate . . . — User learns interactions with the choropleth map; in particular, panning, zooming, reset, and query by location.

4. Interesting trend . . . — User learns that through the visualization, interesting trends may surface . . .

5. Similar design? — . . . and through the details table, the user can read about the history and culture which underlies some of the interesting trends they discover.

6. On your own — Finally, the walkthrough concludes with direction to the settings and about pages, and encouragement for the user to discover new trends!

3.5 Insights Panels

Activated by interactions with the details table, several *Insights* panels are curated to accompany each flag feature. To take an example, the *crosses* insights panel explains how the Nordic cross design is used in the national flags of all independent Nordic countries, and represents Christianity and has an interesting origin in its traditions which date back to 11 June 1748 and regulations regarding Danish merchant ships.

Exemplifying the claimed resuability of the walkthrough system previously described, the insights feature is in fact also built on this system. In particular, the `requirementsSatisfied` field is described by events corresponding to the associated detail queries.

4 Results and Discussion

Many users of our tool had positive feedback. Some were able to discover "surprising insights", and others felt the walkthrough feature was able to "guide the user into searching for their first country". Furthermore, it was some initial user feedback on a very early iteration of this application that the development of some key features in the final interactive visualization; namely, the walkthrough and insights features, as well as the ability to reverse top similar flags, query by dropdown, and properly view trends based on detail queries.

Generally, there might be two kinds of users of the system: A user with limited knowledge of the flags of the world might spend more time absorbing the walkthrough and reading through different historical and cultural knowledge afforded by the insights feature. Then, having gained some familiarity with the system, they might begin to form some simple hypotheses about flag design. Meanwhile, a user, perhaps one with prior interest and experience in vexillology, might be able to more quickly use the application to accomplish some particular goals. They might study particular flag features like stars and crescents, and confirm different hypotheses they might have about these features, or even come up with new insights through the exposed, novel flag similarity calculations.

5 Future Work

In a future version of this project, we would want to experiment with alternative ways of calculating flag similarity. We have discussed a particular set of score component functions, f_B, f_I, and this approach would benefit from a kind of study that could evaluate the effectiveness of the chosen parameters we have described, perhaps compared with alternative parameters. Another idea is to calculate flag similarity based on image processing methods, such as pixel-by-pixel differences.

In curating the insights panels for this project, care and attention was required to select exactly the relevant historical and cultural background, and properly format and display it. A possible extension to this project would be to design a system that can parse the existing, relevant dataset to generate such insights about flag design, or more generally, insights about any arbitrary dataset.

Having built the walkthrough system and other key visualization components in this deomain, we would also want to experiment with visualizing other kinds of datasets, such as how a country's flag has changed over time (see [3]). Another natural extension is to explore the possibility of encoding a country's history as a feature vector, and from there exploring trends of *history similarity*. These extensions would be effective since much of the necessary structure has been built in our work here.

Acknowledgments

The authors wish to thank Arvind Satyanarayan, Katie Bache, Doğa Doğa, Saadiyah Husno, Ebenezer Sefah, for their invaluable guidance and feedback throughout the semester.

References

[1] Flag families, 2013.

[2] List of national flags by design, 2021.

[3] Timeline of national flags, 2021.

[4] M. Dominici. An overview of pandoc. *TUGboat*, 35(1):44–50, 2014.

[5] R. S. Forsyth, E. Moore, and D. Ross. Flags data set, 1986.

[6] R. John. Extensive list of similar flags from around the world, 2019.

[7] D. L. Lechner. Graduate student research instruction: Testing an interactive web-based library tutorial for a health sciences database. *Research Strategies*, 20(4):469–481, 2005. doi: 10.1016/j.resstr.2006.12.017

[8] K. Pletcher. Flags that look alike, 2021.

[9] M. Rahkila and M. Karjalainen. An interactive dsp tutorial on the web. In *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 2253–2256 vol.3, 1997. doi: 10.1109/ICASSP.1997.599500

[10] D. Seyser and M. Zeiller. Scrollytelling – an analysis of visual storytelling in online journalism. In *2018 22nd International Conference Information Visualisation (IV)*, pp. 401–406, 2018. doi: 10.1109/iV.2018.00075