Automatic Interpretation of Map Visualizations with Color-encoded Scalar Values from Bitmap Images

Angela Gabriela Mayhua Quispe

Advisor: PhD. Jorge Poco

Committee Members:

Thesis submitted to the
Department of Computer Science
in partial fulfillment of the requirements for the degree of
Master in Computer Science.

Universidad Católica San Pablo – UCSP
September of 2018 – Arequipa – Peru

To my parents, Santos and Marina, who never stop giving of themselves in countless ways. To my brothers and sister, who encourage and support me.

Abbreviations

CNN Convolutional Neural Network

GIS Geographic Information System

HMM Hidden Markov Model

JSON JavaScript Object Notation

NLP Natural Language Processing

OCR Optical Character Recognition

SVG Scalable Vector Graphics

SVM Support Vector Machine

Acknowledgments

First and foremost, I want to thank God for having guided me throughout these two years of study.

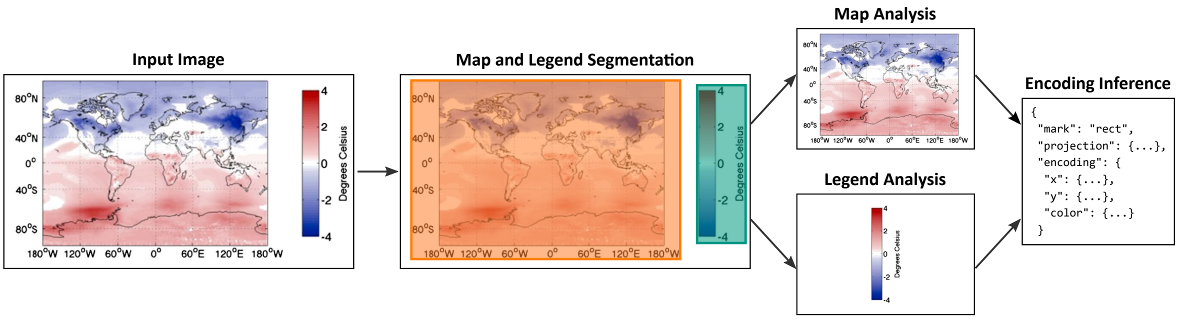
I would like to thank my family for their continuous support and encouragement during these years of study.

I would like to thank in a special way to the National Council for Science, Technology and Technological Innovation (CONCYTEC-PERU) and to the National Fund for Scientific Development, Technological and Technological Innovation (FONDECYT-CIENCIACTIVA), which through the Management Agreement 234-2015-FONDECYT have allowed the grant and financing of my studies in the Master Program in Computer Science at Universidad Católica San Pablo (UCSP).

I would like to express my gratitude and appreciation to my advisor Jorge Poco for giving his guidance and support throughout the Master program and the preparation of this thesis.

Finally, I thank all the people who directly or indirectly helped me in the preparation and presentation of this work, either by exchanging ideas, giving me advice and recommendations or by encouraging me to continue.

Abstract



Map visualizations are used in diverse domains to show geographic data (*e.g.*, climate research, oceanography, business analyses, *etc.*). These visualizations can be found in news articles, scientific papers, and on the Web. However, many map visualizations are available only as bitmap images, hindering machine interpretation of the visualized data for indexing and reuse.

In this work, we propose a pipeline to recover the visual encodings from bitmap images of geographic maps with color-encoded scalar values. We evaluate our results using map images from scientific documents, achieving high accuracy along each step of the pipeline. In addition, we present iGeoMap, our web-based system that uses the extracted visual encoding to enable user-interaction over bitmap images of map visualizations.

Keywords: Visual encoding, Map interpretation, Map visualization.

Resumen

Las visualizaciones de mapas son usadas en diferentes áreas para mostrar datos geográficos (por ejemplo, datos climatológicos u oceanográficos, resultados de análisis empresariales, entre otros). Estas visualizaciones se pueden encontrar en artículos de noticias, artículos científicos y en la Web; sin embargo, muchas de ellas están disponibles como imágenes *bitmap*, lo que dificulta que el computador interprete los datos visualizados para su indexación y reutilización.

En este trabajo proponemos un *pipeline* para recuperar la codificación visual a partir de imágenes *bitmap* de mapas geográficos que utilizan el color para codificar los valores de los datos. Nuestros resultados fueron analizados usando mapas extraídos de documentos científicos, logrando una alta precisión en cada paso del *pipeline*. Adicionalmente presentamos a iGeoMap, nuestro sistema web que utiliza la codificación visual extraída para permitir la interacción del usuario sobre imágenes *bitmap* de visualizaciones de mapas.

Palabras clave: Codificación visual, Interpretación de mapa, Visualización de mapa.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In Section 1.1 we describe the motivation and context of our work, Section 1.2 presents our problem statement. Section 1.3 shows the objectives of this work. Finally, Section 1.5 describes the structure of this thesis document.

1.1 Motivation and Context

Scientific charts are commonly used to visualize quantitative information because they show keypoints and trends among the data. Geographic maps are a popular form of data visualization, used to convey information within a geo-spatial context. The use of maps is not limited to experts such as geographers or cartographers: millions of maps are produced and used by scientists, students, governments, and companies for a variety of analytical purposes (*e.g.*, environmental, economic, political or social). A well-designed map encodes information so as to be interpretable by human viewers; however, these maps are often published as bitmap images, without access to the underlying data. Having access only to pixel-level information impedes automatic processing for tasks such as indexing, search, and analysis (Jung et al., 2017; Siegel et al., 2016) because metadata and pixel values do not include enough information about the content and data plotted on the chart. For that reason, it is difficult to find and reuse map data using either spatial queries (*e.g.*, find all maps involving a specific country) or semantic queries (*e.g.*, find all maps with temperature values in a particular range) (Walter et al., 2013). We need computational solutions to automatically process maps due to the existence of millions of maps that have been digitally scanned or digitally created (Chiang et al., 2014).

Existing methods for automatic chart interpretation focus on analyzing common statistical graphics such as bar, line, area, or pie charts. Some projects attempt to recover the underlying data (Savva et al., 2011; Gao et al., 2012; Al-Zaidy and Giles, 2015; Al-Zaidy et al., 2016; Jung et al., 2017; Siegel et al., 2016; Tummers, 2006), while others focus on recovering the visual encodings (Harper and Agrawala, 2014; Poco and

Heer, 2017). However, these systems do not support analysis of geographic maps. In this work, we extend these prior approaches to recover visual encodings for map images with color-encoded scalar values.

Our primary contribution is a map image analysis pipeline that (i) segments an input image into map and legend regions, and then for each region (ii) identifies text elements, extracts their content using Optical Character Recognition (OCR), and classifies their roles (*e.g.*, legend labels, latitude label, longitude label, *etc.*). Next, our system (iii) determines the type of color legend (*e.g.*, continuous or quantized) and (iv) infers the map projection used (*e.g.*, Equirectangular, Miller, or Robinson). We leverage the extracted text, legend information, and map projection to recover a visual encoding specification in a declarative grammar similar to Vega-Lite (Satyanarayan et al., 2017), a high-level grammar of graphics. An additional contribution is a manually-annotated corpus of geographic map images (each containing a color legend) extracted from scientific papers in the field of climate change, which was used to evaluate our pipeline.

We also present a web-based system named iGeoMap that uses the visual encoding inferred by our pipeline to enable user-interaction over bitmap images of map visualizations. The interactions offered by iGeoMap include recoloring, automatic caption generation, map reprojection and data extraction from map visualization images.

1.2 Problem Statement

Nowadays, there is a large number of geographic map images available on scientific papers, news articles and on the Web; however, users do not have access to the underlying data, and to our knowledge, a method to extract the visual encoding from map visualizations does not exist. For that reason, we propose to apply reverse engineering to map visualizations with color-encoded scalar values and generate as output its corresponding visual encoding. In addition, we developed a web-based system that uses the extracted visual encoding to enable user-interaction from bitmap images of map visualizations.

1.3 Objectives

General Objective

Our main objective is to propose a pipeline to infer the visual encoding in JavaScript Object Notation (JSON) format from a map image with color-encoded scalar values.

Specific Objectives

To achieve our main objective, we have the following specific objectives:

- Extract spatial information from the map plotted on the image.
- Extract color information from the color legend on the image.
- Develop applications that use the extracted information and enable user-interaction.
- Evaluate each step of the pipeline using map images from scientific documents.

1.4 Contributions

This thesis proposes a novel map image analysis pipeline to recover the visual encoding from map visualizations with color-encoded scalar values. A map visualization has two important parts that need to be analyzed: geographic map and color legend. Our contributions are related to each part and are detailed below.

- Extracting and Retargeting Color Mappings from Bitmap Images of Visualizations (Poco et al., 2018).
  - We propose a method to semi-automatically extract color encodings from bitmap visualization images. This color mapping is recovered using color and text information from color legend.
  - We also demonstrate the utility of the proposed method through two user-facing applications: automatic recoloring and interactive overlays.
- Extracting Visual Encodings from Map Chart Images with Color-encoded Scalar Values (Mayhua et al., 2018).
  - Under review in SIBGRAPI 2018.

1.5 Outline

This thesis document is divided into six chapters. After this introduction and problem formulation, in Chapter 2 we survey the literature on map interpretation, automatic chart interpretation and interactive applications from chart images. Chapter 3 presents some basic concepts about the mapping of color and geographic map properties. Next, in Chapter x we describe in detail the corpus, techniques used by our pipeline and their evaluation results. Chapter y presents our web-based system named iGeoMap and its different modules. Finally, the limitations, future works, and conclusions of this work are presented in Chapter z.

Chapter 2

Related Works

Our work draws on prior research in the areas of map interpretation that is focused on extracting information from maps, automatic chart interpretation focused on analyzing charts and interactive applications from chart images that enable user-interaction.

2.1 Map Interpretation

Researchers have proposed various methods to perform automatic *map interpretation* (Walter and Luo, 2011) to extract information from maps and analyze their content. For instance, Dhar and Chanda (Dhar and Chanda, 2006) analyze scanned topographic maps to extract and recognize symbols (*e.g.*, trees, forests, rivers, cities, huts, *etc.*) and text contained within the map. One of the steps is to separate the image into four layers: green elements (trees, forests), red elements (streets), blue elements (rivers, lakes) and black elements (text). The map scale and range of latitude/longitude coordinates are entered by the user to locate points on the map given their geographical coordinates. Finally, the output is an *e-map* that can be used as input to Geographic Information System (GIS). Pezeshk and Tutwiler (Pezeshk and Tutwiler, 2011) also worked on scanned topographic maps; their purpose was to automatically extract each component of the map in separate layers and recognize the text contained. They propose an algorithm for extracting linear features to generate a layer containing map lines (streets, roads, *etc.*); they then use the RANSAC algorithm (Fischler and Bolles, 1981) to improve the text preprocessing and a Hidden Markov Model (HMM) to recognize texts and generate a text output layer.

These previous works focus mainly on topographic maps — *i.e.*, maps characterized by contour lines and road lines (Pezeshk and Tutwiler, 2011) — and recognizing their symbols. Our approach automatically extracts spatial information from the geographical map contained in a map visualization; this information includes the type of geographic projection used by the map and the range of latitude and longitude values in the displayed region.

2.2 Automatic Chart Interpretation

A growing number of techniques focus on the "inverse problem" of data visualization: given a visualization, recover the underlying visual encoding and its corresponding data values (Poco et al., 2018). Some of these approaches have focused on *data extraction*. For instance, ReVision (Savva et al., 2011) classifies images by chart type and extracts data from pie and bar charts to output a relational data table. Similarly, the VIEW system (Gao et al., 2012) extracts information from raster-format charts (*e.g.*, pie, bar, and line charts). It first distinguishes graphical and textual connected-components; depending on the graphic type it applies a different approach to extract data; finally, it generates a data table with that information. Al-Zaidy et al. Al-Zaidy et al. (2016) propose a system that extracts data values from bitmap images of bar charts and generates a semantic graph using the label roles (*e.g.*, x-title, x-labels, y-title, *etc.*); then, the semantic graph is used to generate a summary that describes the input image.

FigureSeer (Siegel et al., 2016) is a framework that extracts information from line charts. It detects the axes to extract their labels and infers their scales through curve fitting. To perform legend analysis, it uses a random-forest classifier (Breiman, 2001) to determine whether or not text serves as a legend label and then obtains its symbol. The analysis of the plotting area is done using Support Vector Machine (SVM) (Cortes and Vapnik, 1995) and a Convolutional Neural Network (CNN) (LeCun and Bengio, 1998) to learn functions and avoid problems caused by the occlusion between the lines. Another application is ChartSense (Jung et al., 2017), an interactive system for data extraction from five types of charts: line, area, radar, bar, and pie charts. Its first step is to classify chart images using a classifier based on GoogLeNet (Szegedy et al., 2015); it then extracts the data using optimized extraction algorithms for each chart type. These approaches extract data from charts that contain discrete legends (*e.g.*, bar, pie, area, line, or radar charts). Our work is focused on the extraction of data from visual components in map visualizations that contain continuous and quantized color legends. This chart type has not been addressed so far, despite being considered in ReVision (Savva et al., 2011) during its classification step.

On the other hand, some methods have been focused on *recovering visual encoding* from a chart. Harper and Agrawala (Harper and Agrawala, 2014) present a tool to decompose and redesign visualizations created with the D3 library (Bostock et al., 2011) (*e.g.*, bar charts, line charts, scatter plots, donut charts, and choropleth). This tool extracts data, marks, and visual encoding by analyzing the Scalable Vector Graphics (SVG) elements of the chart and the data bound to those elements via JavaScript. Poco and Heer (Poco and Heer, 2017) propose a method to recover visual encodings from bitmap images of bar charts, area charts, line charts, and scatter plots; their pipeline identifies textual elements in the image, determines their role within the chart (*e.g.*, chart title, x-labels, x-title, *etc.*), and recovers the text content using OCR. They also trained a CNN (LeCun and Bengio, 1998) for classifying 10 chart types, which achieved an average accuracy better than ReVision and ChartSense, achieving an accuracy of 96% for classifying maps. Using this extracted information they then recover a visual encoding specification. However, their work does not include extraction of color encodings or geographic projections.

As part of this thesis, we presented a work (Poco et al., 2018) where we proposed a technique to extract the color encoding from discrete and continuous legends of chart images, including geographic maps. We identify the colors used and the legend texts, then recover the full color mapping (*i.e.*, associating value labels with their corresponding colors). We continue our thesis work upon that approach focusing on map visualizations; thus, we had to tackle other challenges (such as identifying map projections) and develop new applications enabled by our map image analysis pipeline.

2.3 Interactive Applications from Chart Images

The extracted information from chart images can be useful for different applications. For instance, ReVision (Savva et al., 2011) has an interface to redesign the input chart based on the relational data table extracted by its pipeline. Kong and Agrawala (Kong and Agrawala, 2012) propose to create interactive overlays that are placed above chart bitmap images using the extracted data by ReVision (Savva et al., 2011) and Datathief (Tummers, 2006) pipelines to improve the chart reading.

Kong et al. (Kong et al., 2014) developed an interactive document viewer to improve the reading experience, in this application the user can select a paragraph in a document and some components in the charts are highlighted depending on the selection; they use ReVision (Savva et al., 2011) to extract data from bar charts and a manual annotation interface to recover the original data for other chart types. Other works like ChartSense (Jung et al., 2017) and iVoLVER (Méndez et al., 2016) use semi-automatic approaches to extract data values and also present interactive annotation interfaces to correct the output data and improve the interpretation of charts.

In the same way, we propose a web-based system named iGeoMap that enables the user-interaction on bitmap images of map visualizations. iGeoMap uses the visual encoding generated by our pipeline to create interactive overlays, generate automatic captions, recolor and reproject the input map visualization.

2.4 Final Considerations

This chapter presented some recent proposals related to our thesis work. Some research works have been focused on analyzing topographic maps to extract symbols and texts. On the other hand, other works have focused on extracting data from chart images that contain discrete color legends (*e.g.*, bar charts, line charts, pie charts) to improve the chart understanding through interactive applications.

The next chapter will present some concepts needed to understand better our work, those concepts are related to the mapping of color and geographic map properties.

Chapter 3

Background

In this chapter we present basic concepts that are needed to understand better our work. Section x presents definitions about the mapping of color and Section y details concepts related to geographic map properties.

3.1 Final Considerations

In this chapter, we have presented some concepts about the mapping of color and geographic map properties. Color is used as a visual channel in map visualizations, and the selected colormap type depends mainly on data nature. Furthermore, map projections are used to transform the Earth's 3D surface into a 2D plane.

Next chapter will present our image corpus and how it was created, also presents in detail the steps of our pipeline and their corresponding evaluation results.

Chapter 4

Automatic Interpretation of Map Visualizations

Our work is based on the application of *reverse engineering* to bitmap images of map visualizations to recover their visual encodings in JSON format. To achieve this goal, we extract the spatial and color information from a map visualization image using different techniques, which are presented in this chapter. Section x describes our corpus and how it was generated. Section y presents an overview of our work. Sections v, w and z detail the techniques used to extract spatial and color information; in addition, each one presents the evaluation results. Finally, Section m details how the visual encoding is inferred using the information of previous steps.

Chapter 5

Discussion and Conclusions

The main objective of this work is to recover visual encodings from bitmap images of map visualizations. To achieve our goal, we analyze the map and legend regions using different approaches for each one; after these analyses, we obtain spatial and color information that is used to infer the visual encoding. However, our work has some limitations and possible improvements that are presented in Section 5.1. Finally, our conclusions are presented in Section 5.2.

5.1 Limitations and Future Work

Color legend inside plotting area. As we explain in Section y, we are assuming our map visualizations have the color legend outside the plotting area. This is a common convention in geoscience publications. However, this is not always true in documents from other communities. For example, because of space constraints, authors may try to reduce chart sizes by embedding the legend in the plotting area. As an immediate future work, we plan to explore other techniques from computer vision and propose an algorithm supporting color legends inside the plotting area. Previous works had described some options to tackle this problem (Poco et al., 2018; Siegel et al., 2016); nevertheless, those methods are based on detecting textual information as an initial step, which may propagate more error than alternative vision-based solutions.

Multiple geographic coordinates. There exist multiple standards for geographic coordinates. Our approach assumes that map visualizations contain latitudes and longitudes. However, during our research, we found that other coordinates systems exist such as the *Universal Transverse Mercator (UTM)* and *Military Grid Reference System*. Our technique does not support those systems because we did not find enough map images in the selected geoscience journals; however, we intend to explore new data sources and, if necessary, generalize our techniques to handle those cases.

Maps without coordinates. It is also common to render geographic maps without coordinates, *i.e.*, without textual information indicating latitude and longitude values. Analyzing our image corpus, we did not find many of these cases; however, we know that we can find many of them on the Internet. A possible solution would be to apply techniques from shape analysis and match the map boundaries with pre-built maps in order to identify spatial locations and map projection types.

Discrete legends. Another limitation of our work is the color legend types supported. As we described before, we only support continuous and quantized color legends. An immediate future work would be to extend our work to handle discrete color legends. A partial solution is presented in (Poco et al., 2018); however, that method similarly does not work when the legend lies inside the plotting area.

Improving caption generation. We proposed a single approach for caption generation as a proof of concept, however, more sophisticated techniques can be used. Techniques from Natural Language Processing (NLP) might aid to generate captions with more variability and complexity.

Automatic chart interpretation. The work presented in this thesis is part of a more ambitious objective. The goal is to create an automatic chart interpretation system that, given any chart image, can automatically infer the visual encoding and, as is feasible, extract the underlying data or approximate distributions thereof. If we accomplish this goal, we could create more impactful applications such as improving figure indexing and search, make chart images more accessible for people with disabilities, and perform large-scale analysis of visualization practices.

5.2 Conclusions

In this thesis, we present a novel approach to extract the visual encoding from map visualization images. Given a bitmap map image as input, we generate a visual encoding specification in a format similar to Vega-Lite. We trained and validated each component of our pipeline using real data collected from scientific documents, and our results show high accuracies on each task. Moreover, we developed iGeoMap, a web-based system to improve the map visualization understanding. iGeoMap has different modules which use the output of our pipeline: (1) recoloring to improve perceptual effectiveness, (2) interactive visualizations from a bitmap image, (3) generation of captions for geographic regions based on user's interaction, (4) reprojection to change the design of the map, and (5) recovery of encoded data from map visualizations. In addition, we have created a manually-annotated corpus of geographic map images that were used to evaluate our pipeline and can be used by the scientific community to find and solve new challenges.

Bibliography

Al-Zaidy, R. A., Choudhury, S. R., et al. (2016). Automatic summary generation for scientific data charts. In *AAAI Workshop: Scholarly Big Data*.

Al-Zaidy, R. A. and Giles, C. L. (2015). Automatic extraction of data from bar charts. In *Proceedings of the 8th International Conference on Knowledge Capture*, K-CAP 2015, pages 30:1–30:4. ACM.

Bostock, M. (2017). d3-geo-projection library. https://github.com/d3/d3-geo-projection.

Bostock, M., Ogievetsky, V., et al. (2011). D3: Data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Chiang, Y.-Y., Leyk, S., et al. (2014). A survey of digital map processing techniques. *ACM Computing Surveys*, 47(1):1–44.

Clark, C. and Divvala, S. (2016). PDFFigures 2.0: Mining figures from research papers. In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*, pages 143–152.

Cleveland, W. S. and McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Demir, S., Carberry, S., et al. (2008). Generating textual summaries of bar charts. In *Proceedings of the Fifth International Natural Language Generation Conference*, INLG '08, pages 7–15, Stroudsburg, PA, USA. Association for Computational Linguistics.

Dhar, D. B. and Chanda, B. (2006). Extraction and recognition of geographical features from paper maps. *International Journal of Document Analysis and Recognition*, 8(4):232–245.

Eldersveld, D. (2017). TopoJSON Collection. https://github.com/deldersveld/topojson.

Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395.

Gao, J., Zhou, Y., et al. (2012). VIEW: Visual information extraction widget for improving chart images accessibility. In *19th IEEE International Conference on Image Processing*, pages 2865–2868.

Greenbacker, C. F., Wu, P., et al. (2011). Abstractive summarization of line graphs from popular media. In *Proceedings of the Workshop on Automatic Summarization for Different Genres, Media, and Languages*, WASDGML '11, pages 41–48, Stroudsburg, PA, USA. Association for Computational Linguistics.

Harper, J. and Agrawala, M. (2014). Deconstructing and restyling d3 visualizations. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pages 253–262, New York, NY, USA. ACM.

Hou, X., Yuille, A., et al. (2013). Boundary detection benchmarking: Beyond f-measures. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2123–2130.

Jia, Y., Shelhamer, E., et al. (2014). Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*.

Jung, D., Kim, W., et al. (2017). ChartSense: Interactive data extraction from chart images. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 6706–6717, New York, NY, USA. ACM.

Kong, N. and Agrawala, M. (2012). Graphical overlays: Using layered elements to aid chart reading. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2631–2638.

Kong, N., Hearst, M. A., et al. (2014). Extracting references between text and charts via crowdsourcing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 31–40, New York, NY, USA. ACM.

Krizhevsky, A., Sutskever, I., et al. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, pages 1097–1105.

Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, 7:48–50.

LeCun, Y. and Bengio, Y. (1998). Convolutional networks for images, speech, and time series. In Arbib, M. A., editor, *The Handbook of Brain Theory and Neural Networks*, pages 255–258. MIT Press, Cambridge, MA, USA.

Lucas, S. M., Panaretos, A., et al. (2005). Icdar 2003 robust reading competitions: Entries, results, and future directions. *International Journal of Document Analysis and Recognition*, 7(2-3):105–122.

Mayhua, A., Gómez, E., et al. (2018). Extracting visual encodings from map chart images with color-encoded scalar values.

Méndez, G. G., Nacenta, M. A., et al. (2016). iVoLVER: Interactive visual language for visualization extraction and reconstruction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI'16, pages 4073–4085, New York, NY, USA. ACM.

Munzner, T. (2014). *Visualization Analysis & Design*. K PETERS VISUALIZATION. CRC Press.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.

Pezeshk, A. and Tutwiler, R. L. (2011). Automatic feature extraction and text recognition from scanned topographic maps. *IEEE Transactions on Geoscience and Remote Sensing*, 49(12):5047–5063.

Poco, J. and Heer, J. (2017). Reverse-engineering visualizations: Recovering visual encodings from chart images. *Computer Graphics Forum*, 36(3):353–363.

Poco, J., Mayhua, A., et al. (2018). Extracting and retargeting color mappings from bitmap images of visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):637–646.

Satyanarayan, A., Moritz, D., et al. (2017). Vega-lite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350.

Savva, M., Kong, N., et al. (2011). ReVision: Automated classification, analysis and redesign of chart images. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, pages 393–402. ACM.

Sharma, G., Wu, W., et al. (2005). The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30.

Siegel, N., Horvitz, Z., et al. (2016). *FigureSeer: Parsing Result-Figures in Research Papers*, pages 664–680. Springer International Publishing, Cham.

Smith, R. (2007). An overview of the Tesseract OCR engine. In *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, volume 2, pages 629–633.

Snyder, J. P. and Voxland, P. M. (1989). An album of map projections. Professional Paper 1453, U.S. Geological Survey.

Stokes, M., Fairchild, M. D., et al. (1992). Precision requirements for digital color reproduction. *ACM Transactions on Graphics*, 11(4):406–422.

Szegedy, C., Liu, W., et al. (2015). Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9.

Tajbakhsh, N., Shin, J. Y., et al. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, 35(5):1299–1312.

Tummers, B. (2006). DataThief III. http://datathief.org/.

Turkowski, K. (1990). Filters for common resampling tasks. In Glassner, A. S., editor, *Graphics Gems*, pages 147–165. Academic Press Professional, Inc., San Diego, CA, USA.

Walter, V. and Luo, F. (2011). Automatic interpretation of digital maps. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4):519–528.

Walter, V., Luo, F., et al. (2013). *Automatic Map Retrieval and Map Interpretation in the Internet*, pages 209–221. Springer Berlin Heidelberg, Berlin, Heidelberg.

Whitaker, J. (2016). Matplotlib Basemap Toolkit. https://matplotlib.org/basemap/.

Wu, K., Otoo, E., et al. (2005). Optimizing connected component labeling algorithms. Technical Report LBNL-56864, Lawrence Berkeley National Laboratory.