**Visualizing Uncertainty with Chromatic Aberration**

**Abstract:** In recent years an increasing array of research are being conducted by researchers in the field of uncertainty visualization that attempt to determine the impact of representations on users’ perception and evaluate its effectiveness in decision making. Uncertainties are often an integral part of data and by nature model predictions also contain significant amounts of uncertain information. A prominent example of uncertainty, COVID-19 is a respiratory infectious disease caused by novel coronavirus. Due to its unprecedented challenges over time and frequent changes of strains, scientists and researchers are investigating the available data to discover the patterns in different demographic areas and examine the effect of vaccinations against different variants. In this study, we explore a novel idea for a visualization to present predictive model uncertainties using Chromatic Aberration (CA). We first utilized existing machine learning models to generate predictive results using Covid-19 pandemic data and calculated the corresponding model uncertainties for the most impacted countries with respect to number of new-cases, new-deaths, and new-vaccination for different countries. We then visualized the data itself and its associated uncertainties with an artificially spatially separated channels of red, green, and blue color components. This chromatic aberration representation has been evaluated in a comparative user study. From quantitative analysis it is observed that user is able to identify targets in CA method more accurately than the state-of-the-art VSUP approach. In addition, their speed of target identification was significantly faster in CA as compared to the VSUP method. But their preference between the two does not vary significantly.

**1 Introduction:**Uncertainty visualization is an ongoing area of research but a topic that many practitioners avoid due to the additional complexity that it experiences. There are various studies conducted for uncertainty representations, for example: textual representation such as captions or tooltips [51], graphical representations such as glyphs [21, 54], custom color palettes such as VSUP [35], bivariate choropleth maps [43] and texture patterns [29]. But as far we know, no uncertainty representation has made use of Chromatic Aberration (CA). To accomplish the purpose, we had to go through the steps such as collect relevant data from reputable sources, generate uncertainty information from predictions based data which is accomplished by feeding collected data into machine learning models and calculated uncertainties from the resultant forecasts [6], visualize the uncertainty along with data using CA, as well as competing existing methods in comparable manner, conduct a controlled human-computer interaction experiment to evaluate the effectiveness of the new visual representation, and explain experimental results with numerical analysis and draw conclusions.

In modern world, data is an essential part of human life, and all data has certain amount of uncertainty either in known or unknown form. Since the uncertainties in data can be originated from different phases or sources, it’s important to analyze and measure the amount uncertainty in the data. That’s why we intend to generate uncertainty data from an authentic way. Considering that fact in mind, the study procedure is subdivided into three major phases i. Generate time series forecasted data from COVID-19 data using four machine learning predictive models, ii. Calculate corresponding uncertainties for different countries and visualize uncertainties in terms of Chromatic Aberration (CA) in a graphical presentation, and iii. Conduct user studies to evaluate user perceptions and applicability with commonly used visualizations.

Chromatic aberration is a color distortion or alteration that is sometimes seen on high contrast edges of objects in photographs. Since different colors of light refract to different angles upon traveling through materials with refractive indices [9] (see Figure 1.2), the resulting images may appear to be distorted [10].

**Chart

Description automatically generated**  A picture containing sky, outdoor, building, government building

Description automatically generatedFigure 1.2: Left - different colors of light refract to different angles [10], Right – example of chromatic aberration due to poor quality lens (from wikipedia.com).

CA is an image quality problem so most of the research surrounding CA are conducted to fix the problem and improve image quality thereby. On the other hand, uncertainty is the problem of data quality and relevant research are conducted mostly regarding reducing it to improve data certainty. But existing research conducted to visualize uncertainty is done with traditional approaches such as glyphs. Since our goal is neither to improve image quality nor data quality, we borrowed the term CA for our research to represent uncertainty as a novel approach in the field of visualization.

**2 Related Work**Uncertainty is an unavoidable part of data and due to complexity people usually avoid it to represent in visualization. The term uncertainty can synonymously refer to some other integral property of data such as data quality, error in data, accuracy in prediction, etc. Prediction data generation is becoming way popular day by day by using different machine learning predictive models such as neural network model for instance MLP, LSTM and GRU in [7] for performance evaluation, ARIMA, PROPHET [1, 4, 6] for time series analysis, XGBoost machine learning algorithm for cholera epidemics predictions linked with weather variable [5]. A decision-supporting tool [8] for medical centers and health-care services has been proposed for influenza prediction for Belgium and liver disease predictive analysis in [9]. All these works have been conducted without concerning uncertainty whereas Botchen et al. [29] focuses on uncertainty that occurs during data acquisition and demonstrates texture-based techniques to visualize uncertainty in time-dependent 2D flow fields. Being error in data inherent, improper or eliminated presentations in visualizations can mislead decision making for data analysts, hence Kamal et al. [30] discusses state-of-the-art approaches such as the quantiﬁcation approach to uncertainty visualization, along with the concept of uncertainty and its sources. Bonneau et al. [16] explores uncertainty in the visualization domain by comparing different results (such as a weather forecast) to detect similarities or differences with comparative visualization technique e.g., comparison of border areas than individual pixels.

A statement on the position of uncertainty visualization today is explained in Griethe et al. [18] that defines the basic concept of uncertainty and discusses sources and necessary measures. Through a human-subjects experiment Deitrick et al. [19] integrates visualization with the study of uncertainty, Lundstrom et al. [20] proposes animation methods to convey uncertainty in the rendering and Pang et al. [21] introduces a wide variety of new uncertainty visualization methods like adding glyphs, adding geometry, modifying attributes, modifying geometry, animation and Wittenbrink et al. [54] designs uncertainty vector glyphs, Finger et al. [49] introduces blended icons, Kay et al. [51] presents a novel mobile interface for visualizing uncertainty for transit predictions. Being uncertainty a multi-faceted concept [38, 43] investigates how data uncertainty visualized in maps might influence the process and outcomes of spatial decision-making and determine the accuracy of estimates. The authors Greis et al. [42] published a web-based game on Facebook and compared four representations that communicate different amounts of uncertainty information to the user and compared the results to show how uncertainty regularly influences decision making in our daily life.

From a vision perspective, chromatic aberration leads to various forms of color imperfections or tempering in the image. Koh et. al. [10] presented a user study to observe the effect on users’ judgment with Lateral Chromatic Aberration (LCA) for Chart Reading in Information Visualization on display devices and suggested some guidelines (e.g., using eyeglasses) for designers to avoid such issues and [12, 13] proposed image warping technique to resolve such problems. Colour is widely used in information visualisation to deliver different types of information such as extreme values, patterns and attribute values and color size illusion is one of the common problems in this domain. Yoo et. al. [11] aim to identify appropriate interventions and propose design guidelines of image warping to reduce the illusion effect when size judgement is critical. Real cameras have an aperture through which light falls on an image plane to register the image, but diffraction is an issue on this process, so Lee et al. [15] present a novel rendering system for defocus blur and lens effects by approximating optical aberrations.

**2.1 VSUP**

Both uncertainty visualization and understanding uncertainty are complex and critical tasks. One of the most common approaches of uncertainty visualisation is to encode data values and uncertainty values independently, using two visual variables in a bivariate map. These resulting bivariate maps can be difficult to interpret, and the discriminability of marks can be reduced due to the interference between visual channels. To address this issue, Correl et al. [35] introduces Value-Suppressing Uncertainty Palettes (VSUPs) as in Figure 2.1 (right) whereas a traditional bivariate map shown in Figure 2.1 (left). We highlight this prior work as it is the comparator approach in our user study.

Chart, funnel chart

Description automatically generated  
Figure 2.1: A standard bivariate map (left) and a VSUP (right)

We see that VSUPs allocates smaller ranges of the visual channel to data when uncertainty is high and larger ranges when uncertainty is low. This allocation of visual variables promotes patterns of decision-making that make efficient use of uncertainty information, discouraging comparison of values in unreliable regions of the data, and promoting comparison in regions of high certainty. In traditional bivariate maps Figure 2.1(left), outputs for each combination of value and uncertainty might be represented as a 2D square whereas VSUP approaches it as arcs mapping larger number of outputs for smaller and smaller sets of outputs for higher uncertainty.

But the main limitation of that research is they filter out higher uncertainty values by grouping them altogether which suppresses the values for decision making when uncertainties are high. Due to this higher uncertainty elimination aspect the designers need to carefully consider if this representation is suitable and desirable for certain systems. Another limitation is, since both uncertainty and value are represented by a single color, the perceptual non-separability of color channels are well-known, and which requires the concept of a limited “budget” of distinguishable marks. To achieve the limited budget criteria, it necessitates one to quantize the data. Due to the data quantization, uncertainty visualisation for continuous (or all discrete) values are not possible with limited color budgets.

**Uncertainty Data Generation**

Good quality data is an important part in data visualization research. Without having an authentic dataset research cannot be conducted properly and cannot succeed in the long run. That’s why we chosen COVID dataset from WHO authorized data repository and applied data preparation strategy such as cleaning, validating, and consolidating raw data before feeding to machine learning models. Model are built by exploring and tuning hyperparameters [6, 9] to get better prediction results.

**User Study**

**Results and Numerical Analysis**

**Conclusions and Future Work**

In this thesis, we propose a novel approach for uncertainty visualisation, namely Chromatic Aberration. We conducted a within subject comparative user study with VSUP and our system to assess user performance accuracy/error rate, task completion time, and subjective assessment with NASA-TLX and SUS. From numerical analysis and evaluation of the results, we see user performance and perception is both statistically improved and faster compared to VSUP whereas in the subjective assessment do not vary significantly.

Nevertheless, we note that in real chromatic aberration the chromatic blurring appears continuously from inner edge to outer edge. But in our case, it just gives us a range of uncertainty for the prediction, so the edges are with the same bright color. However, our simplified implementation allows us to reduce the aberration to both double and/or single parameter, which facilitates chromatic aberration tuning with regards to the amount of represented uncertainty. It also allows one to implement the approach relatively easily using standard d3 and SVG operations. However, additional research could be conducted that examine more sophisticated effects. In addition, further research could be conducted with more levels of uncertainties than were tested in both in Correll et al. [35] and the present work, for instance 8-levels instead of 4-levels. The role of CA might also be explored in animated visualizations. And finally, other future work may refine and expand upon some of our other experimental designs such as the starfish streamgraph layout briefly discussed.