**Visualizing Uncertainty with Chromatic Aberration**

By

Md Rashidul Islam  
md313724@dal.ca

Supervised by

Dr. Stephen Brooks

Professor

Faculty of Computer Science, Dalhousie University

Submitted in partial fulfilment of the requirements

for the degree of Master of Computer Science

at

Dalhousie University

Faculty of Computer Science, Dalhousie University

Halifax, Nova Scotia

© Dalhousie University 2020. All rights reserved.

**Table of Contents**

List of Tables ………………………………………………………………………………….. vii

List of Figures …………………………………………………………………………………. viii

List of Algorithms ……………………………………………………………………………. xi

Abstract ……………………………………………………………………………………….. xii

List of abbreviations ……………………………………………………………………...…... xiii

Acknowledgements ………………………………………………………………………….. xiv

**Chapter 1 Introduction** …………………………………………………………………... 1

1.1 Background and Motivation ……………………………………………………….... 1

1.2 Background Concepts & Technologies………..…..………………………………… 2

1.2.1 Machine Learning (Predictive Models) …………………………………... 2

1.2.2 D3.js ……………………………………………………………................. 3

1.2.3 Uncertainty ………………………………………….………...…………… 3

1.2.4 Streamgraph ……………………………………………………………….. 4

1.2.5 Texture …………………………………………………………………….. 5

1.2.6 Chromatic Aberration …………………………….……………………….. 6

1.3 Problem Statement …………………………………………………………………. 7

1.4 Approach ……………………………………………………………………………. 8

1.5 Thesis Outline ……………………………………………………………………….. 9

**Chapter 2 Literature Review** …………………………………………………………… 10

2.1 Introduction ………………………………………………………………………… 10

2.2 Prior works related to prediction in Machine Learning Models …………………… 10

2.3 Uncertainty related prior works ……………………………………………………. 11

2.3.1 VSUP………………………… ……………….…………………………. 19

2.4 Chromatic Aberration related prior works …………………………………………. 20

2.5 Texture related prior works ………………………………………………………… 22

2.6 Evaluation of Visualization Systems ………………………………………………. 24

2.7 Limitations of related works ……………………………………………………….. 26

**Chapter 3 Data Collection, Processing, and Introduction of Models** ……………… 28

3.1 Introduction ……………………………………………………………………….. 28

3.2 Data ……………………………………………………………………………….. 28

3.2.1 Data Collection …………………………………………………………. 28

3.2.2 Sample Data ……………………………………………………………. 30

3.3 Machine Learning Algorithms …………………………………………………… 30

3.3.1 Predictive/Forecasting Models …………………………………………. 31

3.3.2 Time Series Analysis vs Forecasting …………………………………… 31

3.3.3 Concerns of Forecasting ………………………………………………… 32

3.3.4 Example of Forecasting …………………………………………………. 33

3.4 MLP ………………………………………………………………………………... 33

3.5 CNN ……………………………………………………………………………….. 36

3.6 LSTM ……………………………………………………………………………… 38

3.7 ARIMA ……………………………………………………………………………. 40

3.7.1 Auto ARIMA ……………………………………………………………. 41

3.8 Uncertainty Data Generation ……………………………………………………… 42

3.8.1 Uncertainty Data Scaling ………………………………………………... 43

3.8.2 Snapshot of uncertainty data ……………………………………………. 43

3.8.3 Uncertainty Comparison among Models ………………………………… 45

**Chapter 4 Visualization Component Calculations** …..…………………………….... 45

4.1 Introduction ……………………………………………………………………….. 45

4.2 Background Architecture ……………………………………………….………… 45

4.3 Examples of CA in Shapes ……………………………………….………………. 47

4.4 Texture Pattern Generation …………………………………….…………………. 48

4.4.1 Slicing Plot …………………………………………….……………….. 50

4.4.2 Pattern Generation …………………………………….………………… 51

4.4.3 Texture Generation ………………………………….………………….. 52

**Chapter 5 Experimental Designs with Chromatic Aberrations & Texture Patterns**  53

5.1 Introduction …………………………………………………………..……………. 53

5.2 Web Interface ………………………………………………………………..……. 53

5.3 Filtering …………………………………………………………………..……….. 55

5.3.1 Bubble Selection Mode ……………………………….……..………….. 55

5.3.2 Bubble Removal Mode …………………….…………….…..………….. 56

5.4 Legend ……………………………………………………………………..………. 56

5.5 Reshuffling Streamgraph …………………………………………………..………. 57

5.6 Drill-down All Model Predictions …………………………………….….……….. 58

5.7 Star Fish Inspired Design ……………………………………………….…………. 59

5.8 Parallel Coordinate Chart ……………………………………………….…………. 61

5.9 Bubble Grid Chart …………………………………………………….…………… 62

5.10 Horizontal Chart …………………………………………………….……………. 63

5.11 Square Grid Chart …………………………………………………….………….. 65

5.12 World Map ………………………………………………………………………. 66

5.13 Summary of Experiemental Designs ……………………………….…………….. 67

**Chapter 6 Evaluation: User Study Design** …….…………………….…………..……. 68

6.1 Introduction ……………………………………………………….…………………68

6.2 Background and Goal ………………………………………….…………………… 68

6.3 Research Questions ……………………………………………………………….... 69

6.4 Study Material …………………………………………………………………….... 69

6.4.1 Technology and Browser Support………...…………………………….... 69

6.4.2 Study Components………...……………………………............................ 70

6.4.3 Counter Balancing………...……………………………............................ 70

6.5 Recruitment ……………………………………………………………………….... 73

6.5.1 Criteria ………………………………………………………………….... 73

6.5.1 Hiring Procedure ……………………………………………………….... 74

6.5.3 Making Schedule ……………………………………………………….... 74

6.6 Study Procedure .………………………………………………………………….... 74

6.6.1 Start Event …….………………………………….………………….….... 75

6.6.2 Briefing ….…….……………………………….……………………….... 75

6.6.3 Color Blindness Test ….…….…………………….…………………….... 75

6.6.4 Pre-Session Discussion ……..………………….……………………….... 76

6.6.5 Overview of the Questionnaire Structure ………..……….…………….... 76

6.6.6 Component Questions ……………….… .………..…………………….... 79

6.6.7 Example PSQ Questions ……………… ..………..…………………….... 83

6.7 Data Collection and Storing …….…….……….………………………………….... 85

6.8 Session Ending …………...…….…….………….………………….…………….... 86

**Chapter 7 Evaluation: Results and Numerical Analysis**……………………………... 87

7.1 Introduction ………………………………………………..……………………….. 87

7.1.1 Sample Population Demographics ……………..…..…………………….. 87

7.2 Study Results ………………………..…………………..………………………….. 88

7.2.1 Quantitative Questionnaire Results …………..……….…………………..88

7.2.1.1 One-way repeated measures ANOVA ……..………….………..88

7.2.1.2 Paired t-test ………………………………..……………………92

7.2.2 Time Utilization Results …………………………..….…………………. 94

7.2.3 SUS Results …………………………………..……….……………….... 96

7.2.4 NASA-TLX Results ………………………..………….………………....98

7.3 User Comments …………………..…………………..………………………….... 100

7.4 Summary of the Results ……………..……………..………………………………100

**Chapter 8 Conclusion and Future Work** …………………..……….……………… 102

**References** ………………………….…………………………..………………….………101

Appendix A REB Application Form ………………………...……………………………..109

Appendix B Informed Consent Form …………….………...………………………………125

Appendix C First Email …………….…………………..………..…………………………128

Appendix D Second Email …………….……………….……..…….………………………129

Appendix E Examples and Questionnaire ………………..…….………………………….130

E.1 Questionnaire Setup and Arrangement …………….………..…………………130

E.2 Example of CA + Bubble ………………………………………………………134

E.3 Questionnaire on CA + Bubble ………………………………………………...135

E.4 Example of VSUP + Bubble …………………………………………………...137

E.5 Questionnaire on VSUP + Bubble ……………………………………………..138

E.6 Example of CA + Grid …………………………………………………………140

E.7 Questionnaire on CA + Grid …………………………………………………...141

E.8 Example of VSUP + Grid ……………………………………………………...143

E.9 Questionnaire on VSUP + Grid ………………………………………………..144

E.10 Questions on System Usability Scale (SUS) ……………………………….146

E.11 Questions on NASA TLX ………………………………………………….147

Appendix F Ishihara Color Blindness Test Plates ……………….………………………...148

Appendix G Amazon Gift Card Sample ………………………….………………………..154

Appendix H REB Approval Letter ……………………………….………………………..155

Appendix I User Study Results …………………………….….………………………….156

**List of Tables**

Table 3.1 COVID Data property list ………………..………………………… 29

Table 3.2 Screenshot of sample data …………...………………………………30

Table 3.3Top uncertainty countries ………………..….………………………44

Table 3.4 Lowest Uncertainty countries ……………..……………………….. 44

Table 3.5 Uncertainty Comparisons of Models ……..…………………………45  
Table 6.1 Task Arrangement of user study …………..………………………...72

Table 7.1 ANOVA Data Summary ……………………..…………………..… 88

Table 7.2 Shapiro-Wilk Normality Test for ANOVA ……..……….…………..89

Table 7.3 ANOVA Test Results Summary ………………..………….………..90

Table 7.4 Summary of CA vs VSUP performance……...…………………….. 92 Table 7.5 Shapiro-Wilk Normality Test for t-test ………..…..……....……….. 92 Table 7.6 Summary of CA vs VSUP timing ……………..…..…….………..... 94 Table 7.7 SUS scores summary of CA vs VSUP ………..…..…….………….. 98 Table 7.8 Normality test results of NASA-TLX score ….…..…….………….. 99 Table 7.9 Kruskal-Wallis test results of NASA-TLX ….....……….………….100

Table I.1: Questionnaire Raw Scores of four components …..…..………….……156

Table I.2: Questionnaire Raw Scores of CA vs VSUP ………..……………....…157

Table I.3: SUS Raw Scores for CA ……………………..…….…...…………..…158

Table I.4: SUS Raw Scores for VSUP ………………..……….…..…………..…159

Table I.5: NASA-TLX Raw Scores for CA ………..……….…………..……..…160

Table I.6: NASA-TLX Raw Scores for VSUP ………..…….…….….....……..…161

Table I.7 Time Utilization for Full Questionnaire ………..…………..………… 162

**List of Figures**

Figure 1.1 Streamgraph Prototype …………………………….………………… 5

Figure 1.2 Example of Chromatic Aberration ………………………………….. 6

Figure 2.1 Bivariate Map and VSUP …. ………………………………………. 19

Figure 3.1 Predictive modelling workflow to generate uncertainty …...……… 31

Figure 3.2 Example of daily covid forecasting for 200 days ………..….…….. 32

Figure 3.3 Basic Architecture of MLP network …………..…………….……... 33

Figure 3.4 Basic Architecture of CNN network …………..…………….……... 36

Figure 3.5 Basic Architecture of LSTM network …………..…………….……. 39

Figure 4.1 Underlying Geometry of Chromatic Aberration ……..……….…… 46

Figure 4.2 Example CA on Bubbles and Rectangles …………..……………… 48

Figure 4.3 Streamgraph Color Filled (left), Texture Filled (right) ……..………49

Figure 4.4 Sliced Streamgraph …………………………..…………….………..50

Figure 5.1 Initial Web Interface ……………………..……………….…………53

Figure 5.2 Filter by selected countries of interest ………..………….………… 56

Figure 5.3 Removal of countries of interest ………………..………….………..56

Figure 5.4 Reshuffling Main Streamgraph ………………..………….…………57

Figure 5.5 Streamgraphs for all models for Brazil ………..………….…………58

Figure 5.6 Uncertainty presentation on stream by texture …..………….……….59

Figure 5.7 Multi Country Stream Graphs filled by color ……..………………...60

Figure 5.8 Multi Country Stream Graphs filled by Texture of CA ….………… 61

Figure 5.9 Parallel coordinates chart ……………………………….………….. 62

Figure 5.10 Bubble Grid chart with CA textures ……………….….……………. 63

Figure 5.11 Horizontal chart (Color filled) ………………….…….…………….. 64

Figure 5.12 Horizontal chart (Texture filled) ……………….…………………... 64

Figure 5.13 Charts of Daily Counts for different countries …….………………. 65

Figure 5.14 Uncertainty in World view ……………………….………………... 66

Figure 5.15 Zoomed World Map centering Nigeria …………….……………… 67

Figure 6.1 Latin Square (left), Balanced Latin Square (right) ……..………….. 71

Figure 6.2 Example Color Plate in Portal ………………………..……………. 76

Figure 6.3 Flow of Questionnaires for a participant ………………..…….……. 77

Figure 6.4 Module Start View ……..………...………………..………….….… 77

Figure 6.5 Layout of Questionnaire view ……..………...……..………….....… 78

Figure 6.6 Sample Question ……..……...…...…………..……………….….… 78

Figure 6.7 Question-Answer Identification Procedure for CA+Bubble ..........… 80

Figure 6.8 Question-Answer Identification Procedure for CA+Grid ..............… 81

Figure 6.9 Question-Answer Identification Procedure for VSUP+Bubble ...….. 82

Figure 6.10 Question-Answer Identification Procedure for VSUP+Grid ............. 82

Figure 6.11 System Usability Scale Question Example………...………...……... 83

Figure 6.12 NASA-TLX Work-Load Scale Question Example…...…...…...…… 83

Figure 6.13 Session Ending Greetings…...…………………..…………….…..… 86

Figure 7.1 Box plot of user performance …...…….……..……………….…..… 89

Figure 7.2 Normal Distributions for components ...……..……………….…..… 90

Figure 7.3 ANOVA Results...…………………………..………..……….…..… 92

Figure 7.4 Normal Distributions CA vs VSUP...…...…..………..…………..… 93

Figure 7.5 Paired t-test gaussian plot for CA vs VSUP...……..………….......… 94

Figure 7.6 Paired t-test gaussian plot for time utilization...….....…..….……..… 96

Figure 7.7 SUS rating plots for visualization methods ………..………….……. 98

Figure E.1 Balanced Latin Squares ………………………..…………….……. 130

Figure E.2 Questionnaire Email Screen ……………………..…………….….. 131

Figure E.3 Layout of Questionnaire View ……………..……………….…….. 132

Figure E.4 Module Starter View…………………………..……………….….. 132

Figure E.5 Sample Question …………………………………..……….……... 133

Figure E.6 Question-Answer Identification on CA + Bubble …..……..……... 134

Figure E.7 CA + Bubble Questionnaire UI ………………………..…..……… 135

Figure E.8 Questions on CA + Bubble ……………………………………….. 136

Figure E.9 Question-Answer Identification on VSUP + Bubble ……..…..…... 137

Figure E.10 VSUP + Bubble Questionnaire UI ………………………..…..…... 138

Figure E.11 Questions on VSUP + Bubble …………………………..……..….. 139

Figure E.12 Question-Answer Identification on CA + Grid …………..…..…… 140

Figure E.13 CA + Grid Questionnaire UI ……………………………..…..…… 141

Figure E.14 Questions on CA + Grid ………………………………..…...……...142

Figure E.15 Question-Answer Identification on VSUP + Grid ………………... 143

Figure E.16 VSUP + Grid Questionnaire UI ………………….……...………... 144

Figure E.17 Questions on VSUP + Grid …………………………..….….…….. 145

Figure G.1 Amazon Gift Card Sample ………………………….…………….. 154

**List of Algorithms**

Algorithm 3.1 MLP Model ………………………………………………..…. 36

Algorithm 3.2 CNN Model ………………………………………….……….. 38

Algorithm 3.3 LSTM Model ………………………………………….……….40

Algorithm 3.4 ARIMA Model ………………………………………………... 42

Algorithm 3.5 Calculate uncertainty …………………………………………..42

Algorithm 3.6 Data scaling ……………………………………………………43

Algorithm 4.1 CA Construction Formula ……………………………………..47

Algorithm 4.2 Pattern Generation ……………………………………………..52

Algorithm 4.3 Texture Generation ………………………………………………52