



UNIVERSITY OF WEST ATTICA

SCHOOL OF ENGINEERING

**DEPARTMENT OF SURVEYING AND GEOINFORMATICS
ENGINEERING**

Diploma Thesis

**Searching for salient objects on cartographic backgrounds used in online map
services: An experimental study based on eye movement analysis**

Dimitrios Liaskos

Registration number: 17070

Supervisor:

Vassilios Krassanakis

Assistant Professor, UNIWA

Egaleo, July 2023



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΤΟΠΟΓΡΑΦΙΑΣ ΚΑΙ
ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ

Διπλωματική Εργασία

**Αναζητώντας προεξέχοντα αντικείμενα σε χαρτογραφικά υπόβαθρα
διαδικτυακών χαρτογραφικών υπηρεσιών: Εκπόνηση πειραματικής μελέτης
μέσω ανάλυσης οφθαλμικών κινήσεων**

Δημήτριος Λιάσκος

ΑΜ: 17070

Επιβλέπων:

Βασίλειος Κρασανάκης

Επίκουρος Καθηγητής, ΠΑΔΑ

Αιγάλεω, Ιούλιος 2023



UNIVERSITY OF WEST ATTICA

SCHOOL OF ENGINEERING

**DEPARTMENT OF SURVEYING AND GEOINFORMATICS
ENGINEERING**

**Searching for salient objects on cartographic backgrounds used in online map services: An
experimental study based on eye movement analysis**

Μέλη Εξεταστικής Επιτροπής συμπεριλαμβανομένου και του Εισηγητή

Η διπλωματική εργασία εξετάστηκε επιτυχώς από την κάτωθι Εξεταστική Επιτροπή:

A/a	ΟΝΟΜΑ ΕΠΩΝΥΜΟ	ΒΑΘΜΙΔΑ/ΙΔΙΟΤΗΤΑ	ΥΠΟΓΡΑΦΗ
1	Βασίλειος Κρασανάκης	Επίκουρος καθηγητής ΠΑΔΑ	
2	Λουκάς-Μωσής Μισθός	Μεταδιδακτορικός Ερευνητής ΠΑΔΑ	
3	Paweł Cybulski	Assistant Professor Adam Mickiewicz University Poznan	

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος Λιάσκος Δημήτριος του Γρηγορίου, με αριθμό μητρώου 17070 φοιτητής του Πανεπιστημίου Δυτικής Αττικής της Σχολής Μηχανικών του Τμήματος Μηχανικών Τοπογραφίας και Γεωπληροφορικής, δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο Δηλών

Δημήτριος Λιάσκος



Copyright © Με επιφύλαξη παντός δικαιώματος. All rights reserved.

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ και Δημήτριος Λιάσκος,
Ιούλιος, 2023.**

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τους συγγραφείς.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα του και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις θέσεις του επιβλέποντος, της επιτροπής εξέτασης ή τις επίσημες θέσεις του Τμήματος και του Ιδρύματος.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Β. Κρασανάκη, Επίκουρο Καθηγητή του Τμήματος Μηχανικών Τοπογραφίας και Γεωπληροφορικής του ΠΑΔΑ, για την ανάθεση του θέματος της διπλωματικής εργασίας, για την καθοδήγηση και τη γνώση που μου προσέφερε σε όλη τη διάρκεια της εκπόνησης της εργασίας, για την εμπιστοσύνη που μου έδειξε να χρησιμοποιήσω τον εξοπλισμό του εργαστηρίου και για την ευκαιρία που μου έδωσε να πραγματοποιήσω την παρούσα εμπειρική μελέτη.

Θα ήθελα να ευχαριστήσω την οικογένεια μου που με στηρίζει καθ' όλη τη διάρκεια των σπουδών μου και με τη βοήθεια τους κατάφερα να ολοκληρώσω τις σπουδές μου.

Ευχαριστώ όσους εθελοντικά συμμετείχαν στην πειραματική διαδικασία.

Table of contents

Table of figures	11
Table of tables	11
Περίληψη	13
Abstract	15
Chapter 1 : Introduction	17
Chapter 2 : Map design and perception	20
2.1 Definition of cartographic design and the importance of web maps in modern cartography	20
2.1.1 Evolution of cartography in the digital era	20
2.1.2 Principles of cartographic design	22
2.1.3 Usability and interaction in Web Maps	23
2.1.4 Challenges and opportunities	25
2.2 Visual attention and perception, and eye movement analyses on map reading	27
2.2.1 Visual Attention in Map Reading	27
2.2.2 Visual saliency	29
2.2.3 Eye Movement Patterns in Map Reading	31
2.2.4 Cognitive Load and Map Reading	32
2.3 Key findings	33
Chapter 3 : Eye tracking methods	35
3.1 Eye tracking methods and techniques	35
3.1.1 Video-based P-CR eye tracking	35
3.1.2 DPI - Dual-Purkinje Imaging	36
3.1.3 Scleral search coils	38
3.1.4 EOG – Electrooculography	39
3.1.5 Limbus tracking	40
3.1.6 Piezoelectric eye tracking	41
3.1.7 Retinal image-based eye tracking	42
3.2 Fundamental and derived metrics	44
3.3 Environmental conditions	45
Chapter 4 : Experimental study design and implementation	47

4.1 Research design and approach.....	47
4.2 Experimental Design	47
4.2.1 Description of the eye tracking system.....	47
4.2.3 How the eye tracking system works.....	50
4.2.4 Specifications	52
4.2.5 Experimental Stimuli.....	52
4.2.6 Participants.....	54
4.3 Experimental Programming	55
4.4 Recording phase	57
4.4.1 Environmental conditions	57
4.4.2 Experimental procedure	57
4.5 Data Analysis.....	60
4.5.1 Participants' data.....	60
4.5.2 Fixation and saccade indices	65
4.5.3 Grayscale statistical heatmaps generation	66
Chapter 5 : Results.....	69
5.1 Indices analysis	69
5.1.1 Median fixation duration.....	70
5.1.2 Median number of fixations.....	71
5.1.3 Median saccade amplitude.....	73
5.2 Grayscale statistical heatmaps analysis	74
5.3 Grayscale statistical heatmaps comparison.....	93
Chapter 6 : Conclusion and future work.....	106
6.1 Conclusion	106
6.2 Future perspectives.....	107
Chapter 7 : References	109
7.1 Bibliography.....	109
7.2 Web sources	121
Appendix A	122
A.1 Fixation duration indices.....	122
A.2 Total number of fixations indices.....	123

Table of contents

A.3 Saccade amplitude indices	124
Appendix B	126
B.1 Heatmap generator	126
B.2 Heatmaps overlaid on visual stimuli	127
Appendix C	128
C.1 Heatmap comparison.....	128

Table of figures

Figure 3.1. Dual Purkinje Image eye tracker (source: Oculomotor Lab, Department of Ophthalmology and Visual Science, University of British Columbia).	37
Figure 3.2. Scleral search coils (source: CHRONOS VISION GmbH, SSC 2D/3D).	39
Figure 3.3. EOG (source: Metrovision).	40
Figure 3.4. Limbus tracking (source: Jochen Triesch, UC San Diego, http://cogsci.ucsd.edu/~triesch).	41
Figure 3.5. Piezoelectric sensor mounted to the Eyelink II helmet (source: Michael B. McCamy et al., 2013).	41
Figure 3.6. Schematic layout of the binocular modification to the Tracking Scanning Laser Ophthalmoscope. Mirror FM3 splits the field to left and right eye. The result is a split field image with left and right retinal images side by side in each frame. Eyes are not shown to scale (source: Stevenson et al., 2016).	44
Figure 4.1. Eye tracking system (host PC on the left, display PC on the right and the desktop mount eye tracker beneath).	48
Figure 4.2. SR Experiment Builder user interface (source: SR Research Ltd, 2020)	49
Figure 4.3. Data Viewer user interface.	50
Figure 4.4. Output image parameters regarding the resolution, height and width.	53
Figure 4.5. Cartographic backgrounds in the QGIS environment. The red points are the boundaries for the export image and the attribute table contains their coordinates in WGS 84'.	54
Figure 4.6. Gender of the participants.	54
Figure 4.7. Participants' dominant eye.	55
Figure 4.8. Participants' age.	55
Figure 4.9. Structure of the main experiment in SR Experiment Builder UI.	56
Figure 4.10. Visual stimuli used in demo experiment. Cartographic background excerpts from Google Maps.	58
Figure 4.11. Eye tracking system geometry.	58
Figure 4.12. Location of the drift correction target and the chosen color for the targets of calibration, validation and drift correction (RGB [255 0 0]).	59
Figure 4.13. Calibration targets: position, size, color and background color.	60
Figure 4.14. Fixations (cyan) and saccades (yellow) in the EyeLink® Data Viewer UI.	61
Figure 4.15. Fixations (cyan), saccades (yellow), blink saccades (red) overlaid on a random visual stimulus in the EyeLink® Data Viewer UI.	61
Figure 4.16. Excerpt from the gaze data report exported from EyeLink® Data Viewer in a .txt file.	62
Figure 4.17. Excerpt from the saccade report exported from EyeLink® Data Viewer.	63
Figure 4.18. Excerpt from the fixation report exported from EyeLink® Data Viewer.	63
Figure 4.19. Data filters in the EyeLink® Data Viewer UI.	64
Figure 4.20. Excerpt from the string array of the calculated indices.	65

Figure 5.1. Fixation duration, number of fixations and saccade amplitude medians for all visual stimuli, normalized in the range from 0 to 1 with the logarithmic method, visualized in a radar diagram.....	69
Figure 5.2. Bing median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).....	70
Figure 5.3. ESRI median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).....	70
Figure 5.4. OSM median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).....	71
Figure 5.5. Google median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).....	71
Figure 5.6. Wikimedia median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).....	71
Figure 5.7. ESRI median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).	72
Figure 5.8. Bing median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).	72
Figure 5.9. Google median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).	72
Figure 5.10. OSM median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).	72
Figure 5.11. Wikimedia median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).	72
Figure 5.12. Bing median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).....	73
Figure 5.13. ESRI median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).....	73
Figure 5.14. Google median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).....	73
Figure 5.15. Wikimedia median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).	74
Figure 5.16. OSM median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).....	74

Table of tables

Table 4.1. Eye tracker specifications (source: Eyelink 1000 plus user manual).	52
Table 4.2. Cities depicted in the cartographic backgrounds.	53
Table 5.1. Generated grayscale statistical heatmaps with heatmap generator script implementing EyeMMV toolbox (left), overlaid cartographic backgrounds with generated grayscale statistical heatmaps (middle), heatmaps generated from Eyelink Data Viewer (right).	80
Table 5.2. Grayscale statistical heatmap comparisons for all visual stimuli. Every city has four comparison heatmaps with regard to the online map services that the visual stimuli derive from.	94

Περίληψη

Η τρέχουσα γνώση μας για τις αντιληπτικές και γνωσιακές διεργασίες που συμπεριλαμβάνονται στην ανάγνωση χαρτών παραμένει ανεπαρκής. Ο χάρτης λειτουργεί ως μια συμβολική απεικόνιση διάφορων απόψεων του φυσικού κόσμου και των αντίστοιχων χαρακτηριστικών του. Παρά την αφηρημένη φύση του, ένας χάρτης μπορεί ακόμη να είναι ένα οπτικά περίπλοκο ερέθισμα. Όταν μελετούμε φαινόμενα που σχετίζονται με τους χάρτες, η πειραματική έρευνα βασίζεται σε καθιερωμένες τεχνικές από την ψυχολογία και τη νευροεπιστήμη, τα πεδία που ασχολούνται με τη μελέτη της ανθρώπινης συμπεριφοράς και της λειτουργίας του ανθρώπινου εγκεφάλου. Η παρούσα εμπειρική μελέτη επικεντρώνεται στην εξέταση και μοντελοποίηση πτυχών της οπτικής αντίληψης που σχετίζονται με τα χαρτογραφικά υπόβαθρα που χρησιμοποιούνται σε δημοφιλείς διαδικτυακές υπηρεσίες. Συγκεκριμένα, η έρευνα περιλαμβάνει τον σχεδιασμό και τη διεξαγωγή ενός χαρτογραφικού πειράματος που ενσωματώνει τη χρήση τεχνολογίας παρακολούθησης οφθαλμικών κινήσεων για την καταγραφή και ανάλυση τους. Ο κύριος στόχος του πειράματος είναι ο προσδιορισμός χαρακτηριστικών τιμών για βασικούς δείκτες που χρησιμοποιούνται στην ανάλυση των κινήσεων των ματιών, καθώς και η ανάδειξη προεξεχόντων περιοχών ενδιαφέροντος των εξεταζόμενων οπτικών ερεθισμάτων (χαρτών). Οι δείκτες προέρχονται από δύο βασικούς τύπους οφθαλμικών κινήσεων: τις προσηλώσεις και τις σακκαδικές κινήσεις. Επιπλέον, τα δεδομένα που συλλέγονται κατά την πειραματική διαδικασία χρησιμοποιούνται για τη δημιουργία στατιστικών θερμικών χαρτών, οι οποίοι παρέχουν μια συνολική οπτικοποίηση των προτύπων συμπεριφοράς προσοχής αναδεικνύοντας ταυτόχρονα τις περιοχές ενδιαφέροντος της οπτικής προσοχής. Το χαρτογραφικό πείραμα περιλαμβάνει 30 συμμετέχοντες, στους οποίους ζητείται να παρατηρήσουν ελεύθερα 75 οπτικές σκηνές που περιλαμβάνουν χαρτογραφικά υπόβαθρα από πέντε δημοφιλείς διαδικτυακές υπηρεσίες: Google Maps, Wikimedia, Bing Maps, ESRI και OpenStreetMap. Οι οπτικές σκηνές αντιστοιχούν σε χαρτογραφικά υπόβαθρα που αναπαριστούν τρία διαφορετικά επίπεδα μεγέθυνσης: 12z, 14z και 16z. Επιπλέον, για κάθε επίπεδο μεγέθυνσης, επιλέγονται πέντε διαφορετικές ελληνικές πόλεις, με αποτέλεσμα να περιλαμβάνονται συνολικά 15 διαφορετικές πόλεις στις οπτικές σκηνές. Είναι σημαντικό να σημειωθεί ότι οι συμμετέχοντες δεν λαμβάνουν μέρος σε κάποια συγκεκριμένη εργασία χάρτη και απλώς τους ζητείται να παρατηρήσουν τις οπτικές σκηνές. Ο στόχος του προσδιορισμού τυπικών δεικτών για την περιγραφή της οπτικής αντίληψης σε χαρτογραφικά υπόβαθρα, μαζί με την εφαρμογή τεχνικών για την αξιολόγηση της συνολικής συμπεριφοράς κατά τη διάρκεια της παρατήρησης, επιτρέπει τη σύγκριση μεταξύ διαφορετικών οπτικών σκηνών. Αναλύοντας αυτούς τους δείκτες και τις συμπεριφορές παρατήρησης, αποκτούνται πληροφορίες σχετικά με τα μοναδικά

χαρακτηριστικά και προκλήσεις που συνδέονται με τη διαδικασία ανάγνωσης χαρτών. Επιπλέον, η μοντελοποίηση της οπτικής αντίληψης χρησιμοποιώντας δεδομένα που προέρχονται από πειραματικές μελέτες και η χρήση αντικειμενικών μεθοδολογιών μπορούν να παρέχουν πολύτιμες πληροφορίες για τα συστήματα τεχνητής νοημοσύνης. Αυτές οι πληροφορίες στοχεύουν στην πρόβλεψη των αποκρίσεων των αναγνωστών χαρτών. Αυτή η πτυχή είναι κρίσιμη για δύο κύριους σκοπούς: Πρώτον, για την αντικειμενική αξιολόγηση διαφορετικών επιλογών για την χαρτογραφική αναπαράσταση, επιτρέποντας ενημερωμένες συγκρίσεις, και δεύτερον, για την ανάπτυξη χαρτών προσαρμοσμένων στις ανάγκες και τις προτιμήσεις των χρηστών. Οι δείκτες που υπολογίστηκαν παρουσιάζουν ένα ομοιόμορφο αποτέλεσμα, υποδεικνύοντας ότι η οπτική παρατήρηση απαιτεί το ίδιο επίπεδο προσοχής και γνωσιακού φόρτου για κάθε οπτικό ερέθισμα. Ωστόσο, οι στατιστικοί θερμικοί χάρτες δείχνουν ότι τα μοτίβα συμπεριφοράς τείνουν να διαφέρουν, όχι μόνο για κάθε επίπεδο μεγέθυνσης, αλλά και για κάθε ξεχωριστή διαδικτυακή υπηρεσία χαρτών. Συγκρίνοντας τα προεξέχοντα αντικείμενα της ίδιας γεωγραφικής περιοχής, φαίνεται να διαφοροποιούνται σε κάθε οπτικό ερέθισμα.

Λέξεις – κλειδιά: Χαρτογραφία, οπτική αντίληψη, χαρτογραφικά υπόβαθρα διαδικτυακών χαρτών, καταγραφή και ανάλυση οφθαλμικών κινήσεων, στατιστικοί θερμικοί χάρτες.

Abstract

Our current understanding of the perceptual and cognitive processes involved in map reading remains incomplete. A map serves as a symbolic depiction of various aspects of the physical world and its corresponding characteristics. Despite its abstract nature, a map can still be a visually complex stimulus. When investigating map-related phenomena, experimental research draws upon established techniques from psychology and neuroscience, fields dedicated to studying human behavior and the workings of the human brain. The current empirical study focuses on examining and modeling aspects of visual perception related to the cartographic backgrounds utilized in popular online services. More specifically, the research involves designing and conducting a cartographic experiment that incorporates the use of eye tracking technology for recording and analyzing eye movements. The experiment's main goal is to identify characteristic values for key indices used in eye movement analysis, as well as to highlight salient areas of interest of the examined visual stimuli (maps). These indices are derived from two fundamental types of eye movements: fixations and saccades. Furthermore, the data collected in the experiment is utilized to generate cumulative grayscale statistical heatmaps, which provide an overall visualization of attentional behavior patterns, simultaneously highlighting the areas of interest of visual attention. The cartographic experiment engaged a group of 30 participants who were asked to freely observe 75 visual scenes featuring cartographic backgrounds sourced from five popular online services: Google Maps, Wikimedia, Bing Maps, ESRI, and OpenStreetMap. The visual scenes encompassed map backgrounds representing three distinct zoom levels: 12z, 14z, and 16z. Additionally, for each zoom level, five different Greek cities were selected, resulting in a total of 15 unique cities included in the visual scenes. It's important to note that the participants were not given any specific tasks and were simply instructed to observe the scenes. The objective of determining standard indices for describing the visual perception on cartographic backgrounds, along with the application of techniques to assess overall observational behaviors, enables comparisons to be made between different visual scenes. Analyzing these indices and observational behaviors, insights can be gained regarding the unique characteristics and challenges associated with the map reading process. Furthermore, modeling visual perception using data obtained from experimental studies and employing objective methodologies can provide valuable inputs for artificial intelligence systems. These inputs aim to predict the behavioral responses of map readers. This aspect is crucial for two main purposes: Firstly, for the objective evaluation of different options for cartographic representation, allowing for informed comparisons and secondly, for the

development of user-oriented maps tailored to specific needs and preferences. The indices calculated present a uniform result, indicating that the visual exploration requires the same amount of attention and cognitive load for each visual stimulus. However, grayscale statistical heatmaps have proven that the behavioral patterns tend to differ, not only for every zoom level, but also for each distinct online map service too. Comparing the prominent features of the same geographical area for each service, salient objects differentiate in each visual stimulus.

Key words: Cartography, visual perception, online map cartographic backgrounds, eye movement tracking and analysis, grayscale statistical heatmaps.

Chapter 1 : Introduction

Nowadays, online map services have become prevalent in our daily lives. Indeed, the ubiquity of digital technologies has made accessing maps and spatial information an integral part of our lives (Liu et al., 2020). From navigating to planning trips, these digital platforms provide users with access to thorough cartographic backgrounds, facilitating spatial understanding and decision-making. However, their effectiveness heavily relies on the user's ability to locate and identify the information needed, such as prominent landmarks, points of interest and navigational signals (Keskin et al., 2023). Understanding how users visually process and interpret the objects on such cartographic backgrounds is critical for improving the design and usability (Roth et al., 2011).

Eye tracking stands out as a highly valuable method for evaluating various aspects associated with the interaction between a map and its user. Since maps correspond to visual stimuli, examination during the observer's reaction is needed (Ciolkosz-Styk, 2012). Cartographic studies using eye movement analysis have been performed since 1971 (Steinke, 1987) (reported in Krassanakis & Cybulski, 2021) with various conclusions referring to the way visual attention is directed depending on the visual variables used. Since Steinke's (1987) initial comprehensive examination and review of eye movement studies in cartography, the impact of studies using the eye tracking method is becoming increasingly evident in contemporary research.

Effectiveness and efficiency of the elements of the cartographic symbolization align with the speed and accuracy of detection of the map user (Garlandini & Fabrikant, 2009). Extensive research has been conducted on how these terms are affected by examining the visual patterns during the search of symbols on cartographic backgrounds or by using different visual variables (Garlandini & Fabrikant, 2009) (reported in Krassanakis & Cybulski, 2021). Moreover, as also reported in Krassanakis & Cybulski (2021), there have been empirical studies testing the usability of digital map products, comparing visual behavior on paper and digital maps (Incoul et al., 2015), examining the effectiveness and efficiency of interactive (Çöltekin et al., 2009) and animated maps (Dong et al., 2014).

Map designers strive to create visually appealing and informative maps, but it is important to examine at the same time how users actually perceive and most importantly prioritize the depicted objects of a map (Duchowski, 2007). The insights gained from past research represent another stride towards comprehension of the mechanisms underlying visual attention on cartographic products and the visibility and prominence of salient objects. Krassanakis et al., 2013, in a previous study examining

saliency models, demonstrated limitations in accurately predicting human visual behavior on topographic maps.

The present work is included in the wider object of cartography and more specifically is referred to map visual perception. In the context of the diploma thesis, an empirical study is designed to investigate the ground truth of visual attention on cartographic backgrounds used in online map services, employing eye movement analysis to obtain insights into how salient objects are detected. This empirical study is investigating map reading process and the efficiency and effectiveness of the cartographic symbolization elements. More specifically, the study focuses on how these elements affect map reading process by conducting an eye tracking experiment, utilizing contemporary methods of eye movement tracking and analysis. This diploma thesis is organized in seven chapters and three appendices. The **first** chapter is the introduction.

The **second** chapter refers to map design and perception, describing how cartographic data visualization defines the effectiveness and efficiency of the cartographic backgrounds used in online map services. Map exploration requires visual attention from the user, hence there is a review on theories of eye movement analysis in the context of map reading, saliency models and key findings useful for future perspectives, that the present thesis is fulfilling.

The **third** chapter contains distinct eye tracking methods, as well as the fundamental and derived metrics in eye tracking. Moreover, the environmental conditions during an experimental process are described, in order to distinguish the optimal conditions.

The **fourth** chapter refers to materials and methods, respecting the experiment design and the approach of this experimental study. It describes various notorious eye tracking methods, including the one used in this study, and how the eye tracking system being used works. Additionally, there is a narration of how the experiment is conducted to achieve data collection, including the external factors affecting the experiment, and consequently perform data analysis using the EyeMMV toolbox. The data analysis is divided in three subchapters describing how the participants' data are stored and manipulated, in order to compute indices to analyze how participants explore the visual stimuli and generate grayscale statistical heatmaps based on binocular raw gaze data, revealing the salient locations. Comparisons between grayscale statistical heatmaps, regarding the zoom level or the service the visual stimulus is derived from, assist to explore quantitative and qualitative differences.

The analysis is followed by the **fifth** chapter which contains the results. The indices calculated are examined thoroughly to understand the visual behavior of the participants on the visual stimuli, the grayscale statistical heatmaps are analyzed to extract quantitative and qualitative interpretations of

the salient objects explored and comparisons are made to enrich the insights gained on quantitative differences.

The **sixth** chapter concludes the findings of the experimental study and lists suggestions for future perspectives.

The **seventh** chapter contains the references and websites used for the conduction of this experimental study.

The **appendix A** contains the scripts for the computation of the indices regarding the fixation and saccade reports.

The **appendix B** contains the scripts for the grayscale statistical heatmaps generation and the overlaying of the cartographic backgrounds with the grayscale statistical heatmaps.

The **appendix C** contains the script for the grayscale statistical heatmaps comparison.

Chapter 2 : Map design and perception

2.1 Definition of cartographic design and the importance of web maps in modern cartography

2.1.1 Evolution of cartography in the digital era

In the field of cartography, the map has served not only as a product but also as a subject of scientific investigation for a considerable period (Eckert, 1921). The strategic plan of the International Cartographic Association (ICA) puts forth a definition of the map concept that has evolved over time. Currently, the map is defined as follows: “*A map is a symbolic representation of geographical reality, depicting selected features or characteristics that result from the creative decisions made by its creator, and it is intended for use when spatial relationships are of primary importance*” (<http://icaci.org/strategic-plan>).

Numerous definitions of a map can be found, including the one presented in The History of Cartography Encyclopedia by Woodward & Lewis (1998), who define maps as “*graphic representations that aid in comprehending the spatial aspects of objects, concepts, conditions, processes, or events in the human world.*” It is noteworthy that this definition places emphasis on the outcome of map utilization: enhancing human understanding of reality. Additionally, Robinson & Petchenik (1976) offer a concise definition describing a map as “*a visual portrayal of the environment.*”

In our dynamic and constantly evolving world, the concept of cartography and its primary output, the map, have experienced significant transformations, reflecting the broader spatial turn in society. The traditional perspective of the map as a representation crafted by a cartographer, aiming to visually convey a comprehensive and mostly unchanging database of geographic information to the user, has given way to a more participatory and collaborative approach. The map has transcended its role as a static window to the world and has evolved into an interactive, mobile, dynamic, and collaborative interface. It now serves as a bridge between individuals, groups of people, and the ever-changing environment. This shift acknowledges the active involvement of users in contributing to and shaping the map, blurring the boundaries between map creators and map users. The map has become a platform for engagement, allowing users to interact with and manipulate spatial information, fostering a more immersive and personalized experience. This transformation in cartography reflects the growing recognition that maps should be adaptable, responsive, and capable of capturing the

evolving nature of our surroundings. It also aligns with the increasing emphasis on collaboration, co-creation, and collective intelligence in spatial representation. Consequently, the map has become a powerful tool for facilitating communication, decision-making, and understanding within a society that embraces a more participatory and dynamic approach to spatial information.

The collection of Web 2.0 tools and techniques that enable the creation and advancement of dynamic maps on the internet are well-known as web mapping technologies. These technologies, along with the associated maps and geographic information contributed online, form what is commonly known as the GeoWeb. The term GeoWeb is used to describe the body of geographic information and the corresponding maps that are shared and accessed through web mapping technologies (Leclerc et al., 2001; Haklay et al., 2008). As reported in Roth et al. (2014), numerous recent works have emerged that delve into the social and ethical results of volunteered contributions to the GeoWeb, shedding light on its significance and timeliness. Crampton, (2009), Elwood (2010), Goodchild (2007), Harvey (2012), Wilson (2012) and Sack (2013) have provided valuable insights into this subject matter, exploring various aspects and implications associated with the volunteered contributions made to the GeoWeb.

Present-day web mapping technologies, along with web technologies in general, can be classified into three main categories. The first category includes server-side technologies used to index and query geographic information. This can be done either from a centralized source or from distributed sources, such as cloud-based platforms. The second category includes client-side technologies that are responsible for manipulating web maps within the user's web browser. Lastly, the third category comprises web services or intermediary scripts that facilitate the communication of information requests between the client and server. It is crucial to carefully design and develop all three of these technologies to ensure a web map's architectural effectiveness. This approach to technological development is vital for creating a well-functioning web mapping system. The details of these categorizations are discussed by Roth et al. (2008).

In the context of web mapping, it is argued that a clear distinction can be made between the roles of a cartographer and a GIS technician, primarily due to their divergent contributions to client-side technology (Roth et al., 2014). The cartographer's involvement lies in the client implementation, which encompasses design considerations related to the map representation. This includes decisions regarding the choice of projection, generalization, symbolization, typography, and other visual aspects of the map. The cartographer also plays a role in shaping the user interface provided for manipulating the map and aims to create an optimal user experience with the web map. In contrast, the GIS

technician's focus typically revolves around the server-side technologies and data management aspects of web mapping. Therefore, the increasing significance of client-side technology has led to greater involvement and contribution from cartographers in shaping the visual and interactive aspects of web maps.

Throughout the history of cartography, societal needs and technological advancements have significantly shaped the field. One notable example is the impact of Google Map Service, since 2005, as they have provided widespread access to maps through their user-friendly and interactive online map interface. This innovation has facilitated everyday tasks such as locating places, finding routes, and navigation, making maps more accessible and useful to the public. The convergence of societal needs and technological innovations continues to drive the evolution of cartography.

Online map services commonly utilize the web Mercator projection, which was introduced by Google Maps and is based on the WGS84 reference ellipsoid. The scale of the map is determined by the chosen zoom level selected by the user and is represented by an integer number ranging from 0 (smallest scale) to 18 (largest scale), with each zoom level change altering the map scale by a factor of 2. This dynamic scaling enables the display of different scale maps for the same area. The appropriate level of generalization on a map is influenced by its purpose and scale. When comparing a specific area on online map services at the same zoom level but from different services, variations in the degree of generalization and information density can be observed (Skopeliti & Stamou, 2019).

2.1.2 Principles of cartographic design

Cartographic symbolization plays a crucial role in the map production process as it involves the use of specific design principles to visualize qualitative and quantitative variations in geographic data. For static maps, these principles align with the visual variables introduced by Bertin (1983), while animated maps employ dynamic variables as proposed by DiBiase et al. (1992) to fulfill their function. In addition to visual and dynamic variables, acoustic variables as suggested by Krygier (1994) can also be incorporated to support the cartographic design process. In recent years, eye tracking studies have yielded valuable insights into the fundamental role of design principles (Alaçam & Dalc, 2009; Brychtova & Coltekin, 2016; Cybulski & Krassanakis, 2022; Garlandini & Fabrikant, 2009; Krassanakis, 2009). The findings from these studies can directly inform the selection of cartographic symbolization for more effective map design.

The selection of the appropriate cartographic representation, including the map type and its corresponding symbology, is determined by the characteristics of the data to be mapped. The

guidelines for creating such cartographic representations have been proposed by Bertin (1967). Bertin suggests a systematic analysis of the data, considering factors such as whether it is discrete or continuous, qualitative, or quantitative, and so on, and linking these data attributes with the perceptual properties of map symbols and their variations. Although Bertin's approach was primarily based on intuition, many of his ideas have started to be empirically validated (Garlandini & Fabrikant, 2009) and find support from advancements in other scientific fields, such as psychology, vision science, and cognitive science.

The process of abstraction plays a vital role in designing an effective map. It involves removing irrelevant details from the map display to ensure clear communication of the represented geographic phenomenon. By abstracting unnecessary information, the map can convey the essential elements in a concise and understandable manner (A. H. Robinson, 1995). Generalization, which involves reducing data complexity, can have significant effects on line length, local coordinate density, topology, and other properties. To preserve these characteristics, various algorithms have been developed (Cromley & Campbell, 1992; Buttenfield, 2002). Additionally, there is a body of literature focused on database support for managing multi-scale data and enabling flexible zooming across multiple scales and map purposes. Notable studies in this area include the work of Jones et al. (2000).

Visualization serves as a collaborative medium for humans and computers to leverage their respective capabilities and achieve optimal outcomes. Spatiotemporal phenomena and processes exist and operate across different spatial and temporal extents. For instance, a hailstorm is a localized and short-term phenomenon, while climate change spans the globe and unfolds over extended periods (G. Andrienko et al., 2010). The scale of spatial analysis is determined by the size of units used to measure phenomena and aggregate measurements. It is widely recognized in geography that the scale of analysis can significantly influence results. Patterns or relationships observed at one scale may not be evident at another, and in extreme cases, different scales can reveal contradictory relationships. Scale variance is thus an important property associated with geographic features.

2.1.3 Usability and interaction in Web Maps

Presently, web maps have evolved to become adaptive, meaning they adjust based on usage and user context (*e.g.*, Reichenbacher, 2003; Friedmannová et al., 2006). They are interactive, allowing users to make requests and receive real-time responses (*e.g.*, G. L. Andrienko & Andrienko, 1999; Roth, 2013). Additionally, web maps have become mobile-friendly, tracking the user's location as they navigate through the depicted area (*e.g.*, Clarke, 2004; Meng & Reichenbacher, 2005). They are also

designed to be multiscale, offering the ability to view the world at various geographic locations and extents (*e.g.*, Brewer & Buttenfield, 2007; Roth et al., 2011). Furthermore, web maps can be updated in real-time, reflecting geographic events and processes as they occur (*e.g.*, Boulos & Burden, 2007; Goldsberry, 2007) (as cited in Roth et al., 2014).

Visual analytics diverges from conventional analytical approaches by leveraging interactive visual representations to enhance human abilities in pattern recognition, establishing connections, and making inferences. However, this assumption requires empirical validation (Fabrikant & Lobben, 2009). Geovisualization is increasingly employed in experiments to test this assumption and serves as an analytical method for recording collected data. Some experiments combine accuracy measurements, information-seeking time, and eye tracking methods to assess the efficiency and effectiveness of visual displays in decision-making contexts.

Interactive visual interfaces hold immense potential in facilitating the empirical exploration of suitable scales of analysis and verifying results through scale modifications and different aggregation methods. To fully realize this potential, more understanding is needed regarding appropriate visual representations for distinct types of data at various spatial and temporal scales. Corresponding analysis techniques supporting interaction should enable seamless transitions and comparisons among scales and aggregations. While progress has been made in this direction, further research is necessary (Laube et al., 2010; Wood, 2005) (as cited in G. Andrienko et al., 2010). The study of scale issues in visual analytics can draw upon advancements in cartographic generalization, including theories, best practices, and algorithms for geometric and semantic generalization of diverse data types (Mackaness & Ruas, 2007).

One of the primary strengths of well-crafted maps is their ability to visually communicate existing knowledge and facts effectively. With online mapping technology, these maps can be disseminated globally, raising awareness about the current situation in various domains. By presenting information spatially, maps can provide a comprehensive and easily understandable overview of a particular subject. Maps serve as powerful tools for representing the present situation, allowing users to comprehend complex information briefly. They can depict a wide range of data, such as demographics, land use, infrastructure, environmental factors, and more, enabling a holistic understanding of a given area or phenomenon. Whether it is displaying the distribution of populations, the extent of natural resources, or the impact of a particular event, maps facilitate the visualization and analysis of current conditions. Besides, maps can track and document changes over time. By creating successive maps at different points in time, one can observe and analyze the evolution, growth, or

decline of various phenomena. This temporal dimension adds a valuable perspective to understanding trends and patterns of our world. By documenting changes through maps, we can identify shifts in demographics, land cover, urbanization, climate, and other factors, contributing to a deeper comprehension of the past, present, and potential future scenarios. Overall, well-crafted maps, especially when combined with online mapping technology, play a vital role in increasing awareness and disseminating knowledge about the current situation.

Decision-makers at diverse levels can monitor and compare indicators for policy development and action across different geographical scales by utilizing interactive map dashboards linked to geographic databases. These dashboards incorporate space-time analytical capabilities and operate at multiple scales, enabling decision-makers to make informed choices and take appropriate actions based on the data presented.

2.1.4 Challenges and opportunities

The rapid advancement of technology in web mapping presents both advantages and disadvantages for web cartographers. The ongoing progress in technological solutions brings about a fundamental conflict: as web mapping options become more versatile and extensive, it becomes increasingly challenging to stay proficient in teaching and utilizing these technologies effectively. The origins of web mapping can be traced back to the early days of the internet. It is commonly associated with the emergence of the World Wide Web in 1991 or the introduction of the first graphical user interface browser in 1993 (Peterson, 2008). In the initial stages of web mapping, the focus was on sharing digitally scanned, static maps (Cartwright, 2008).

In the past, geographic data and map production were primarily accessible to professional cartographers who had the necessary training. However, there has been a shift in recent times with the increased availability of (open) data and the emergence of user-friendly GIS and mapping software online. This has enabled non-trained individuals with internet access to create their own maps. This increased accessibility has led to a significant rise in the number of maps being created today. However, the quality of these maps in terms of effective communication is not always convincing. To address this, cartography has drawn inspiration from developments in human factors and human-computer interaction research, and advancements in personalization technology.

The focus of many scientific research projects in cartography now lies in the use and usability of maps. The development of new cartographic products involves user requirement studies that engage users from the initial stages of the design process. By considering the information needed and

questions of the users, map display solutions can be tailored to their specific requirements. Furthermore, with the increasing involvement of the crowd in map-making projects, the user and the mapmaker often become the same person.

One of the strengths of cartography lies in its ability to adapt to societal and technological changes without compromising its core values. It is crucial to continually develop robust cartographic theory supported by empirical evaluations to effectively utilize the available data sources and leverage the advancements in technology both now and in the future. Over the past decade, advancements in web-based mapping have expanded the quantity and diversity of maps accessed by the public daily. The widespread availability of web maps has created numerous opportunities for Cartography, including the potential for highly interactive, location-aware, on-demand, and real-time maps.

Inexperienced map creators often rely on the pre-defined color schemes offered by commercial mapping and GIS software without realizing that they have the option to customize them. They may not be aware of the potential impact of poorly chosen color schemes on the map's readability and accuracy. Default color schemes provided by GIS software are often unappealing and lack guidance on their effective utilization. It is crucial to recognize that these default schemes may not be suitable for all types of maps and mapping purposes. If color schemes are not thoughtfully constructed and applied to the data, readers can become frustrated, confused, or even misinformed by the map.

In some cases, changes to map display and geometry are interconnected, meaning that modifying one can affect subsequent modifications to the other. However, previous research has emphasized changes to geometry over display changes in multi-scale mapping, resulting in an increased workload for map creators. Cartographic managers faced with balancing or optimizing the workload must make decisions such as selecting the appropriate data resolution for specific display scales, determining when to change map symbols during a multi-scale progression, identifying when data geometry should be modified for smaller scale presentations, and deciding when new data compilations are necessary (Brewer & Buttenfield, 2007). Symbol modification is often less resource-intensive than geometry modification, making it a viable strategy for reducing the overall workload for map designers. As geographic phenomena interact across multiple scales, analytical tools must adequately support cross-scale analyses. The research must address the question of how analysts, encompassing a range of disciplines, can uncover and comprehend cross-scale relationships between phenomena.

2.2 Visual attention and perception, and eye movement analyses on map reading

2.2.1 Visual Attention in Map Reading

Given the evident necessity to investigate how map users react when observing realistic cartographic products such as static maps, animated maps, and interactive multimedia maps, a considerable number of experimental studies have been conducted by the cartographic community to explore various aspects of the map reading process. The process of reading maps is a multifaceted visual procedure that can be influenced by several factors related to both cognitive and perceptual aspects of human vision. To comprehend how maps function, it is necessary to conduct experimental studies that investigate the visual behavior and strategies employed by map users. As maps can be considered visual stimuli, it is reasonable to apply research from related fields to the study of cartography. Therefore, there is a clear need to examine how map users react while observing distinct types of cartographic products.

Investigating visual behavior involves the conduction of experiments that include visual stimuli of diverse forms, one of them being cartographic backgrounds. Comprehension of visual behavior in the context of map reading comes from analyzing and visualizing the experimental gaze data (Krassanakis, 2021). Usability is defined by ISO 9241-11 and eye tracking is an unbiased method for usability studies, because the information acquired is not self-reported (Dolezalova & Popelka, 2016). Fixation and saccades are fundamental metrics in eye tracking. Metrics based on fixation and saccade events are calculated in eye movement analysis, helping gain more insights. The analysis is aided by various visualization methods that help with the understanding of patterns and dynamics of the eye movements.

Eye tracking is a particularly valuable method for evaluating several aspects related to the process of reading maps. While early studies examining eye movements in cartography date back to 1971, with researchers like Williams investigating symbol selectivity using geometric symbols, and Jenks pioneering the analysis of eye movements to understand visual attention in dot maps in 1973, the past decade has witnessed a growing interest in the cartographic community regarding the application of eye tracking techniques. This interest is evident in the increasing number of studies published in recent years. Additionally, several workshops and training sessions have been organized by the community to promote knowledge and understanding in this field, including the ICA workshop "Eye tracking: why, when, and how?" in 2013 (<https://icaci.org/invitation-to-workshop-eye-tracking->

[why-when-and-how](#)), the series of Eye Tracking for Spatial Research (ET4S) workshops (https://et4s.ethz.ch/2020/?page_id=53), and the ICA workshop on "Designing and conducting user studies" (<https://cogvis.icaci.org>).

Steinke's (1987) comprehensive review marked the initial exploration of eye movement studies in cartography and related fields, such as psychology. In our days, the impact of eye tracking methods is readily apparent in cartographic research. A recent study by (Kiefer et al., 2017) offers a well-documented overview of the valuable contributions made by eye tracking technology in spatial research, encompassing spatial cognition, geographic information science, and cartography. According to Kiefer et al. (2017), contemporary eye tracking studies in cartography build upon theoretical concepts developed in research domains that investigate human visual behavior, particularly in the context of "map reading" conditions (as cited in Krassanakis & Cybulski, 2019).

For several decades, psychologists and researchers from other related fields have been documenting and studying the eye movements of individuals to gain insights into their reading patterns and visual exploration of static graphic displays. These displays include texts, advertisements, artworks, charts, and diagrams. Scholars such as Wade & Tatler (2005) and Duchowski (2007) have contributed to this field by conducting research and providing valuable insights into eye movement analysis.

Ever since the pioneering eye tracking studies conducted on visual displays (Buswell, 1935); (Yarbus & Yarbus, 1967), it has been established that individuals tend to direct their fixations towards visually captivating and informative areas within a scene. This phenomenon, commonly observed in eye movement research, demonstrates that people exhibit a natural inclination to concentrate their visual attention on scene regions that are engaging and likely to provide relevant information (Henderson & Ferreira, 2004).

The study of eye-movement behavior allows researchers to determine the areas of a visual display that people focus their attention on. By examining the relationship between attentional focus and the information content present at those locations, it becomes possible to gain deeper insights into the design considerations of a graphic display used for problem-solving. Researchers such as Grant & Spivey (2003) and Thomas & Lleras (2007) have utilized this correlation to evaluate design aspects in a more targeted manner.

Having the ability to objectively measure and study a map user's visual behavior, along with the visual strategies employed when observing distinct types of visual stimuli, is a fact that makes eye tracking an invaluable tool. Eye tracking techniques have been successfully applied in various domains to investigate the process of reading maps. These applications include exploring the design principles

and functions of cartographic symbolization, comparing map users' perception when observing 2D and 3D representations, and assessing the usability of interactive cartographic environments such as Web GIS platforms (as cited in Krassanakis & Cybulski, 2019).

Examining visual perception while observing symbols with a variety of shapes, poses a significant challenge, as it directly impacts the complexity of maps. To address this challenge, an experimental study conducted by Bargiota et al. (2013) investigated the observation of both theoretical and real cartographic lines with distinct geometry. The study aimed to determine whether critical points, which are essential for line generalization algorithms, were observed by map users. The study results indicated fixations were in regions of the lines containing critical points.

In addition to the primary components of a map, such as cartographic symbols and annotations, secondary elements may also influence the map reading process. Edler et al. (2020) conducted an eye tracking experimental study focused on a memory task, which revealed the impact of map legend position on the performance of map users. Specifically, the results highlighted that map legends positioned on the right side were decoded faster compared to those positioned on the left side.

2.2.2 Visual saliency

Eye movement predictions that rely solely on saliency maps (SMs) are exclusively influenced by preattentive processes and do not consider cognitive factors such as task knowledge, visual display characteristics, or domain expertise (Henderson et al., 2007). Salient regions within an image are commonly recognized as areas that attract greater human attention. The detection of salient objects enhances the efficiency of scene processing by extracting the most significant and informative components of a scene. Most visual search models, whether they involve observable eye movements or hidden shifts of attention, rely on the notion of a saliency map. This saliency map is a specific representation that encompasses a two-dimensional map indicating the level of saliency or prominence of objects present in the visual surroundings.

One approach implemented to salient object detection algorithms is the bottom-up method. Such methods are data-driven and utilize low-level data such as features, colors, and spatial distances to generate saliency maps. Various computational models have been proposed at this level. One such model is the central-surrounded approach introduced by Itti et al. (1998), which extracts image features across multiple scales to detect salient objects. Harel et al. (2006) proposes a graph-based visual saliency method, using the Markov chain to describe dissimilarity. Cheng et al. (2014) compute saliency

values based on global contrast differences and spatial weighted coherence scores. Tong et al. (2014) present a method for salient object detection using multi-scale superpixel segmentation. These bottom-up saliency detection methods offer valuable information without prior knowledge. However, they may struggle when processing images with complex backgrounds.

Only a small portion of the information captured by our visual system actively influences our behavior, despite the shared belief that we perceive everything in our visual environment. Change blindness is the phenomenon that makes this fact clear, where substantial changes in images go unnoticed under normal viewing conditions. However, when specifically directed to those changes, observers have no difficulty perceiving them (Simons & Levin, 1997; O’regan et al., 1999) (as cited in Itti & Koch, 2000).

Both overt and covert attention play a crucial role in regulating access to privileged levels of information and ensuring that the selected information aligns with behavioral priorities and goals. From an operational standpoint, information is considered "attended" when it enters and remains in short-term memory, for a sufficient duration to be consciously reported. Therefore, visual attention is intricately connected to visual awareness, as highlighted by Crick & Koch (1998)

However, the process of selecting a specific spatial location is still a matter of investigation. Is it primarily driven by sensory cues from the environment (bottom-up processing), or does the expectation of target characteristics play a crucial role? A significant body of literature has focused on the psychophysics of visual search, examining how individuals orient themselves towards sparsely positioned targets or targets in natural scenes using either unseen or visible shifts of attention. For comprehensive reviews on this topic, refer to the works of Itti et al. (1998) or the survey article by Toet et al. (1998).

Considerable evidence supports a two-component framework for the control of attention allocation within a visual scene. This framework has been supported by numerous studies (Bergen & Julesz, 1983; Braun & Julesz, 1998; Braun & Sagi, 1990; Nakayama & Mackeben, 1989). The first component biases the observer towards selecting stimuli based on their saliency, a bottom-up mechanism utilizing center-surround mechanisms. The second component is a slower, top-down mechanism influenced by cognitive and volitional factors, which guides attention like a "spotlight" under variable selection criteria. However, there is still ongoing debate regarding whether visual consciousness can be reached solely through saliency-based or top-down attentional selection, or if both mechanisms are involved.

The concept of a saliency map for preattentive selection was introduced by Koch & Ullman (1985). Objects in the visual environment have their saliency encoded on a two-dimensional map. This idea is similar to Treisman's (1998) concept of a "master map." Within this saliency map, there is competition among neurons, resulting in a single winning location that corresponds to the most salient object, which becomes the next target of attention. If this location is subsequently inhibited, the system automatically shifts to the next most salient location, providing internal dynamics to the search process.

Several computational models of human visual search have incorporated the concept of a saliency map in various forms. Researchers such as Treisman (1988), Olshausen et al. (1993), Wolfe, (1994), and Itti et al. (1998) have developed models that utilize a saliency map. The advantage of employing an explicit saliency map is that it provides a straightforward mechanism for integrating input from multiple feature maps that operate independently. This integration process yields a single output: the next location that should be attended to in the visual scene.

Most of the research in this area has focused on analyzing eye movements in response to natural scenes that contain real-world objects, which are typically represented through sketches, line drawings, or photographs. While the Itti et al. (1998) model has shown success in predicting eye movements in simple displays used in attention experiments, its effectiveness in predicting gaze patterns in meaningful scenes has been limited (Henderson, 2003; Torralba et al., 2006). It is important to note that this approach has not been extensively tested or validated for more abstract representations such as maps, diagrams, or other types of graphic displays. However, there have been promising findings regarding the applicability of saliency map-based models in cartographic work, specifically for static and dynamic thematic map displays (Fabrikant & Goldsberry, 2005), as well as for remotely sensed images (Swienty et al., 2007).

2.2.3 Eye Movement Patterns in Map Reading

In recent years, numerous approaches for visualizing gaze data have been suggested, ranging from basic methods like scanpath plots to more advanced techniques that can highlight the visual patterns of multiple observers (Burch et al., 2015). One significant challenge in gaze data visualization is the creation of 2D static visualizations that effectively represent spatiotemporal data (Räihä et al., 2005).

The advancement of eye tracking technology has facilitated the collection of eye movement data, which has consequently increased the demand for improved visual analysis methods (Krygier,

J.B., 2008). Various established techniques exist for visualizing eye tracking data, such as heat maps, focus maps, and gaze plots (scan paths). In eye tracking studies, a significant objective is to comprehend the differences and similarities in eye movements among subjects. Previous research has introduced several approaches to mitigate visual clutter and enable comparisons across multiple participants (*e.g.*, Keates, 2014; Horbiński, 2019; Golebiowska et al., 2020). Experts often annotate areas of interest (AOIs) on stimuli and associate them with fixations to capture the semantics of eye movements. Drawing inspiration from the field of text visualization (*e.g.*, Liao et al., 2019; Popelka et al., 2019), earlier studies have also proposed visualization techniques to support AOI-based analysis (*e.g.*, MacEachren, 2004; Keskin et al., 2020).

A statistical grayscale heatmap is a quantitative output that is generated by utilizing point data distributions, which can be based on either raw gaze data or the centers of fixation points (Krassanakis, 2021). Additionally, grayscale statistical heatmaps can serve as a visualization method for aggregated gaze data, allowing the representation of the overall spatial distribution of visual attention from multiple observers. Several eye tracking datasets (Perrin et al., 2020) include such visualizations with the aim of providing an objective ground truth that can be utilized in deep learning approaches. Furthermore, in a recent study by Cybulski & Krassanakis (2022), grayscale statistical heatmaps were employed for the development and computation of a set of statistical indices that aimed to indicate the visual strategies of map users while performing a specific target-based cartographic task. Specifically, they proposed five indices to examine the impact of map label language on visual search, which were used to dispense map readers' visual attention in the central and peripheral areas of different cartographic backgrounds, while searching for point symbols labeled in Polish and Chinese languages.

2.2.4 Cognitive Load and Map Reading

Eye movements serve as observable indicators of how individuals allocate their attention during visual inference tasks. Cognitive scientists leverage eye tracking data to gain insights into the underlying cognitive processes involved in various visual cognition tasks. By analyzing eye movement records, researchers can infer valuable information about how individuals process visual information with scholars such as Grant and Spivey (2003) and Henderson and Ferreira (2004) utilizing eye tracking methodologies to advance our understanding of cognitive processes in visual cognition.

Moreover, the cognitive operations related to visuo-spatial processing occur exclusively during periods of eye fixations. This characteristic of eye movements makes them particularly valuable for

studying spatial cognition matters in disciplines such as geography, Geosciences, and other related fields. Eye-movement studies offer a compelling approach to delve into the complexities of spatial cognition and explore the intricate relationships between visual attention and the processing of spatial information. Researchers, recognizing this potential, have turned to eye-movement studies to investigate spatial cognition issues (Irwin, 2004).

These studies build upon theories and models established in related research domains such as psychology and neuroscience. The research findings generated over the years can directly inform the design of maps that are more effective and efficient in achieving their intended objectives. In the 1970s, cartographers began using eye-movement recording techniques to investigate how individuals visually engage with static maps (Steinke, 1987). Their aim was to enhance the design of map products by incorporating insights from eye-movement research, thus creating more user-friendly and improved maps (Montello, 2002). However, despite the first interest in eye movement studies on maps during the 1970s and early 1980s, the collection of eye-movement data in cartography significantly declined. Montello (2002) proposed that one factor contributing to this decline was the perceived costliness and technical difficulties associated with eye-movement analysis. Some critics argued that this data collection method did not provide mapmakers with new insights beyond what they already knew, making it less justifiable in terms of effort and expense (Montello, 2002). Another reason for the limited success of eye-movement studies in cartography may be attributed to the focus on determining where individuals looked, rather than delving into the how and why of eye fixation behavior or connecting it to specific tasks (Brodersen, 2003). To gain a deeper understanding of eye fixations on maps and to elucidate why certain cartographic displays are more effective than others, it is suggested that cartographers need to situate their studies within cognitive theory (Fabrikant et al., 2010). This approach would involve interpreting eye fixations in the context of cognitive processes, exploring the underlying mechanisms, and establishing the relationship between eye movements and specific cognitive tasks.

2.3 Key findings

Currently, with most maps distributed as digital products featuring animated and interactive content, the primary challenge in cartographic research is evaluating its usability. In a study conducted by Incoul et al. (2015), the visual behavior of map users was compared while observing both paper and digital cartographic maps. The study revealed that although visual attention allocation was similar in both products, digital maps resulted in a higher frequency of fixations per second.

Numerous studies have also investigated the effectiveness and efficiency of interactive maps (Çöltekin et al., 2009) and animated maps (Opach & Nossun, 2011; Dong et al., 2014), including those distributed within GIS environments (Brychtova et al., 2012) or on the Internet as web maps (Alaçam & Dalc, 2009; Manson et al., 2012). Special visualization techniques such as cartograms (Ullah et al., 2016; Krassanakis, 2021) have also been examined. The results of these studies have revealed significant differences among various products, particularly when considering the map reading process involving the use of various tools and buttons within their graphical user interface (GUI). Furthermore, the function of typical visual variables may vary in dynamic cartographic visualizations. For instance, Dong et al. (2014) found that the role of variables such as color hue and size was related to the display size, which is critical given the diverse range of devices with digital monitors used to view maps today.

Some of the aforementioned studies combine eye tracking methodology with additional experimental techniques such as video recording, and questionnaire analysis to examine usability issues. These studies underscore the essential need to investigate how map readers interact with the map interface during the modern process of map reading.

Moreover, the creation and widespread availability of ground truth eye tracking datasets produced from map reading empirical studies, can contribute to the implementation of artificial intelligence techniques such as machine learning and deep learning. This advancement has the potential to make future maps and map services more intelligent and tailored to users' specific needs under different viewing conditions. The knowledge gained in the past decade represents a significant step towards a deeper understanding of how visual attention operates during map observation. The insights obtained from these studies can directly inform the design of more efficient and effective cartographic products. Eye tracking data can also be utilized to model user behavior while performing specific map tasks, as demonstrated by the modeling approach employed by (Dong et al., 2018).

Chapter 3 : Eye tracking methods

Eye tracking is a research technique that refers to the recording of gaze coordinates (Gökstorp & Breckon, 2022) spatiotemporally, during the observation of visual scenes. As reported in Holmqvist et al. (2021), eye tracking examines eye movements, gaze behavior, and changes in pupil size. Default algorithms and settings of the eye tracking systems can obstruct the studies because environmental factors, setup variations and participant related factors need to be taken in consideration and to be quantified to assess the quality of the acquired data (Holmqvist & Andersson, 2017). Three measures are the main criteria for quality evaluation: accuracy, precision, and data loss. Accuracy is the estimation the eye tracker is able to do of the real location the participant is looking (Nyström et al., 2013). It can be described as the gap between the gaze position recorded and the actual gaze position. Precision is the ability of the tracker to maintain consistency when reproducing a gaze position while the true gaze is constant or unchanged (Niehorster et al., 2020). Data loss is the amount of data missing in the eye tracker signal (Holmqvist & Andersson, 2017). In the end of the experiment, a calibration and a second validation come particularly useful to evaluate if the accuracy and the precision remained in the desired range.

3.1 Eye tracking methods and techniques

3.1.1 Video-based P-CR eye tracking

Merchant (1967) introduced video-based P-CR eye tracking, which has become the dominant method in the market as of 2021. Camera-based P-CR eye trackers utilize the pupil center (P) and one or more reflection centers (CR) from infrared illuminators in the eye tracker. Gaze direction is estimated by calculating the relative positions of the P and CR coordinates within the pixel coordinate system of the video image. This is typically done by subtracting the CR coordinate from the P coordinate. It is worth noting that more sophisticated models have been developed in recent years (Hansen & Ji, 2009).

The combination of pupil and CR signals in the P-CR gaze signal has been found to potentially amplify post-saccadic oscillations and lead to an overestimation of peak saccadic velocity, as noted by Hooge et al. (2016). When discussing potential technologies for future studies on saccade dynamics, Hooge et al. (2016) suggest that CR-tracking variants without involving the pupil could be a preferred method. However, Holmqvist & Blignaut (2020) reported inaccurately measured amplitudes of small

eye movements (below 2 degrees of the visual angle) in all eleven P-CR eye trackers they tested. They attribute this issue to erroneous calculations of the CR center by the eye trackers' image processing algorithms, which interact with the resolution of the eye camera sensor.

Variations in the CR signal can be caused by changes in head position relative to the eye tracker, which can influence the size and shape of corneal reflections, as highlighted by Guestrin & Eizenman (2006). Additionally, factors such as illumination levels, sampling frequency, and the quality of the camera's optic lenses can impact the CR signal. Droege & Paulus (2010) emphasize that the use of low-quality eye cameras can further reduce precision in the gaze signal due to slower pixel updating. This can result in pixels retaining some brightness from the previous corneal reflection, creating a bright trace alongside the actual reflection. Consequently, accurate center calculation of the CR image becomes more challenging and less reliable.

The P-CR method is the one used in this study, but many other methods have been developed to record eye movements with overviews provided showing similarities and differences between them from many studies over the years (Holmqvist & Andersson, 2017; Hansen & Ji, 2010; Duckowski, 2007; Young and Sheena, 1975) (cited Holmqvist et al., 2023).

3.1.2 DPI - Dual-Purkinje Imaging

The Dual-Purkinje Imaging (DPI) system is an analog eye tracker that estimates gaze by analyzing the relative movement of the infrared reflection off the corneal (P1) and the reflection at the back of the crystalline lens (P4). In Figure 3.1, a DPI tracker is presented. The DPI system reports the positions of P1, gaze, and head translation as voltages (Crane & Steele, 1985). The continuous signal produced by the DPI can be digitized at the desired sampling frequency using an AD-converter. The DPI system was widely used in psychology laboratories and featured in influential publications such as Frazier & Rayner (1982) and Deubel & Schneider (1996). However, the DPI system has a higher learning threshold compared to P-CR trackers, and its main drawback is that it is a bulky and delicate machine built with optoelectronics from the 1970s, with commercial servicing available from only one person. Nevertheless, Intoy & Rucci (2020) have developed a camera-based DPI system that maintains data quality comparable to the original analog system while incorporating modern electronics, which has the potential to revive the DPI measurement technique.

In DPI eye tracking, the P1 corresponds to the same reflection as the CR in P-CR trackers. However, there is a crucial difference between the two. P-CR eye trackers estimate the center of the

CR based on a small portion of a pixelated camera image, whereas the DPI system determines the center of an analog light beam. This distinction has been suggested as the reason the DPI system does not inaccurately measure the amplitudes of small eye movements, as reported by Holmqvist and Blignaut (2020).

The DPI eye tracking method is capable of capturing gaze signals with sufficient quality to detect tremor, oculomotor drift, microsaccades, and smooth pursuit reliably. Detailed information about its performance can be found in publications by Holmqvist and Blignaut (2020) and Ko et al. (2016). (Holmqvist, 2016) reported a median precision of 0.008° and an accuracy of 0.4° across 192 participants, surpassing the capabilities of any video-based P-CR system. However, the quality of DPI data tends to be lower when recording participants with small pupils that cover the P4 reflection, leading to inaccuracies and potential data loss (Crane & Steele, 1985; Holmqvist et al., 2020). To obtain optimal DPI recordings, it is recommended to work with participants who have large pupils, either in dark rooms or with artificially dilated pupils. Additionally, it is worth noting that the reliance on the P4 reflection in DPI results in the measurement of the largest amplitudes of post-saccadic oscillations among all eye tracking systems (Deubel & Bridgeman, 1995).



Figure 3.1. Dual Purkinje Image eye tracker (source: Oculomotor Lab, Department of Ophthalmology and Visual Science, University of British Columbia).

3.1.3 Scleral search coils

The technique of scleral search coils was first introduced by (D. A. Robinson, 1963) and later adapted for use with human participants by (Collewijn et al., 1975). This method involves placing a copper wire coil, embedded in an annulus or contact lens, onto the sclera of the eye. By subjecting the participant to oscillating magnetic fields, the induced voltage in the eye coil can be used to determine the orientation of the eye in relation to the magnetic fields. Collewijn (1998) referred to this technique as the "gold standard" of eye tracking. (Reulen & Bakker, 1982) presented the double magnetic induction principle, which was further improved by (Bour et al., 1984). Like the DPI, scleral search coil systems are analog trackers, and the data can be digitized at high sampling frequencies. These coils can record both eye and head rotation for the same participant (Collewijn et al., 1985). In a comparison between a coil system and a torsion-capable video eye tracker, (Houben et al., 2006) found that the gaze signal from the coil system was ten times more precise. Another study by Ko et al. (2016) compared a coil system to a DPI and concluded that while the coil system provided slightly higher precision, both systems offered sufficient data resolution for reliable detection of intersaccadic (fixational) eye movements. Collewijn (2001) sampled data at a frequency of 10,000 Hz and reported a tracking range of 20° in all directions with a resolution of 1'. Malpeli (1998) reported a precision of 1' (0.017°), and Collewijn et al. (1988) recorded saccades with amplitudes of up to 80°.

The scleral coil tracking method is considered to be significantly invasive, and there is evidence suggesting that older coil systems, when combined with the application of anesthesia, could temporarily reduce visual acuity (Irving et al., 2003) and cause deformation of the visual field (Duwaer et al., 1982). However, contemporary search coils have overcome some of these issues by being embedded on flexible contact lenses. These modern coil systems are used in research and clinical diagnostic applications in neuro-ophthalmology and neurology. This is primarily due to their high precision and the fact that patients often experience uncontrolled head and body movements, necessitating a tracking method that can accommodate such movements. Figure 3.2 depicts scleral search coils from CHRONOS VISION.



Figure 3.2. Scleral search coils (source: CHRONOS VISION GmbH, SSC 2D/3D).

3.1.4 EOG – Electrooculography

Recordings of the horizontal component of gaze were achieved by Schott (1922) and Meyers (1929) through the utilization of the corneoretinal potential principle discovered by Du Bois-Raymond in 1849. The electrooculography (EOG) technique captures eye movements by employing electrodes placed on the sides of the eyes, which detect the electromagnetic field generated by the corneo-retinal electrical potential ranging from 10 to 30mV (Brown et al., 2006). This signal is then processed through an isolated instrumentation amplifier and connected to a chart recorder or computer. EOG is an analog method and is commonly integrated into other recording devices. For example, electroencephalogram (EEG) systems often incorporate additional electrodes for eye-related measurements that can be utilized for EOG recordings. In Figure 3.3, the setup of the EOG on the observer is presented.

When investigating larger eye movements, electrooculography (EOG) can be a valuable form of eye tracking. However, smaller movements tend to be overshadowed by the noise present in EOG data. An advantage of EOGs is their ability to be utilized even when the eyes are closed, such as in the study of rapid eye movement (REM) sleep (Aserinsky & Kleitman, 1953). Nevertheless, EOG eye tracking exhibits lower accuracy compared to most other methods. Young and Sheena (1975) reported an average inaccuracy of 1.5°–2° in EOG measurements.



Figure 3.3. EOG (source: Metrovision).

3.1.5 Limbus tracking

The first documented implementation of a limbus tracker, specifically a photo-electric version, was presented by Torok et al. (1951). Limbus trackers function by estimating the boundary between the iris and sclera, utilizing either video or photosensors. Photodiode-based limbus eye trackers were commercially available for research purposes until the year 2000 through the Skalar company. However, their usage now primarily revolves around controlling lasers during eye refractive surgeries (Arba-Mosquera & Aslanides, 2012). It is important to note that the Ober Saccadometer, although not a limbus tracker, operates as a corneal bulge tracker (Holmqvist & Andersson, 2017). Similar to the Skalar limbus tracker, the Saccadometer employs photosensors to monitor the corneal bulge. Video-based limbus trackers leverage the fact that the limbus border, situated between the iris and sclera, exhibits a contrast comparable to that of the pupil-iris border. However, limbus trackers are not susceptible to artifacts stemming from pupil-related issues, a concern for both DPI and P-CR systems. Additionally, refraction in the cornea does not pose a problem for limbus trackers. Eye trackers equipped with low-resolution cameras may benefit from employing the limbus tracking method. Nonetheless, a significant challenge arises when a substantial portion of the limbus becomes obstructed by the eyelid, thereby introducing complications in image processing. The equipment of the limbus tracking is observed in Figure 3.4.

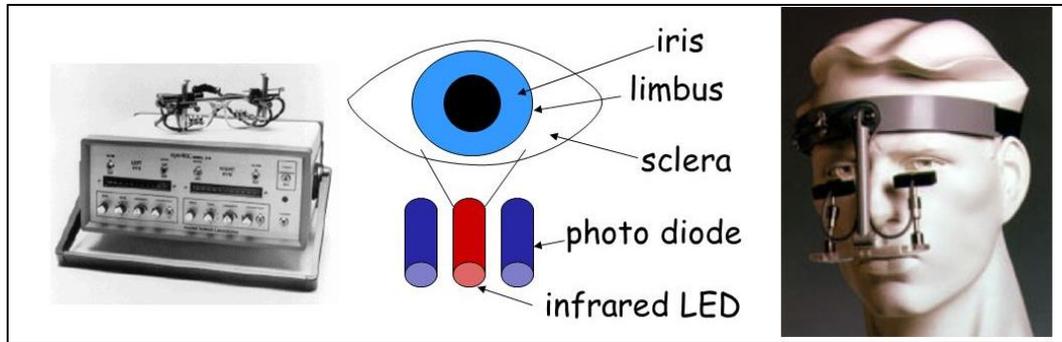


Figure 3.4. Limbus tracking (source: Jochen Triesch, UC San Diego, <http://cogsci.ucsd.edu/~triesch>).

3.1.6 Piezoelectric eye tracking

The piezoelectric transduction method, initially presented by (Bengi & Thomas, 1968), involves the utilization of a silicone tipped piezoelectric bimorph that contacts the sclera, typically positioned in the interpalpebral region near the temporal limbus. This method generates voltage signals capable of detecting horizontal microsaccades and oculomotor tremor. However, this analogue eye tracker has primarily been employed for measuring intrafixational eye movements and has not found application in other areas. There is a concern that the applied pressure on the sclera may influence microsaccade behavior (McCamy et al., 2013). The piezoelectric equipment appears in Figure 3.5.

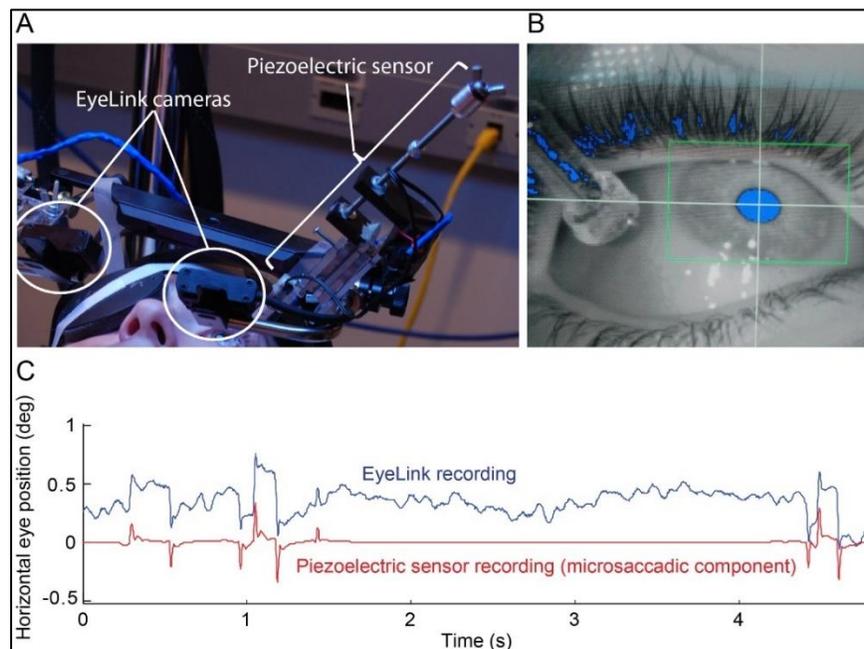


Figure 3.5. Piezoelectric sensor mounted to the EyeLink II helmet (source: Michael B. McCamy et al., 2013).

3.1.7 Retinal image-based eye tracking

The computational tracking of retinal features, initially pioneered by Cornsweet (1958), involves the identification of the optic disk, blood vessels, and smaller features in retinal images. This process utilizes computer vision algorithms to analyze the movement of these features within the camera's field of view and infer eye movements. Retinal image-based eye trackers are considered the most accurate and precise among all existing eye tracking methods. In Cornsweet's early system (1958), although limited to tracking features along a single axis, it was capable of detecting eye movements (microsaccades) with amplitudes as small as 10 seconds of arc (0.0028°). Putnam et al. (2005) achieved remarkable results in terms of gaze position accuracy (5 arcseconds or 0.0014°) based on snapshots captured using an adaptive optics retinal camera.

The most advanced and precise retinal-based eye trackers are typically constructed using scanning imagery, particularly scanning laser ophthalmoscopes (SLO). These trackers utilize the "rolling shutter" principle to capture eye motion, as demonstrated by Mulligan (1997). SLOs equipped with adaptive optics, which provide high resolution, high magnification, and densely sampled retinal video, are particularly effective for this purpose, as highlighted by Stevenson and Romero-Borja et al. (2005). A significant advancement was made by Stevenson et al. (2016) who introduced the first binocular system, which optically divided a single SLO image field between the two eyes. This development expanded the capabilities of retinal-based eye tracking technology. In Figure 3.6, appears the system introduced by Stevenson et al. (2016).

Retinal imaging systems often obstruct the forward view, which poses a limitation on stimulus presentation. However, advancements have been made to address this issue. Bartuzel et al. (2020) describe a microelectromechanical system (MEMS)-based retinal imaging system that allows for simultaneous stimulus presentation and high-frequency recording (1240Hz). Despite this improvement, the trackable range of retinal imaging systems tends to be narrower compared to other types of eye trackers. Bartuzel et al. (2020) report a range of 16 degrees (8 degrees left and 8 degrees right), which can be contrasted with the 20-40 degrees range of DPI and many video-based P-CR trackers, as well as the 90 degrees or greater range of scleral coils.

Until now, the use of retinal-image-based eye trackers has been restricted in terms of their application. The inherent compromise between accuracy and range has primarily made them valuable for investigating eye movements during stable fixation (Bowers et al., 2019). Retinal eye trackers have

primarily found utility in the field of ophthalmology, particularly in studying retinal diseases and their effects on vision and subtle eye movements (Godara et al., 2010).

The various eye tracking technologies mentioned can be configured to capture data from either one eye (monocular) or both eyes (binocular). Binocular eye tracking is commonly used in remote eye trackers, where the signals from the left and right eyes are combined by averaging synchronous data samples in the recording software, often referred to as "cyclopean gaze." A study by Cui and Hondzinski (2006) suggests that averaging the signals from both eyes improves accuracy, although Hooge et al. (2019) found that this was only the case for certain participants. Head-mounted eye trackers can experience parallax errors, which occur when the eye's viewpoint and the scene camera's viewpoint do not align, particularly when the measurements are not confined to a single plane. Binocular averaging is frequently employed in glasses-based eye trackers (such as SMI ETG and Tobii Glasses) and the Ober Saccadometer to mitigate the parallax issue. The geometry of parallax error has been thoroughly investigated by Mardanbegi et al. (2012), Batista Narcizo & Witzner Hansen (2015), and Tatler et al. (2019). Alternatively, the signals from the two eyes can be utilized to measure vergence, as demonstrated by Liversedge et al. (2006). Jaschinski et al. (2010) showed that the EyeLink II, under ideal conditions, can detect vergence eye movements with a depth of just under 40mm at a viewing distance of 60cm. However, vergence measurements with P-CR eye trackers are susceptible to artifacts that can affect accuracy, as reported by Hooge et al. (2019), including the influence of pupil size on vergence. During calibration for binocular recordings, there is a choice to calibrate both eyes simultaneously or separately, as discussed by Nuthmann & Kliegl (2009), and Švede et al. (2015). Furthermore, X. Wang et al. (2019) found that the calculation of the vergence point (the intersection between the gaze direction vectors of the left and right eyes) can exhibit a significant deviation from the fixated point, with a wide distribution in depth and a misestimation of the mean vergence point toward the participant.

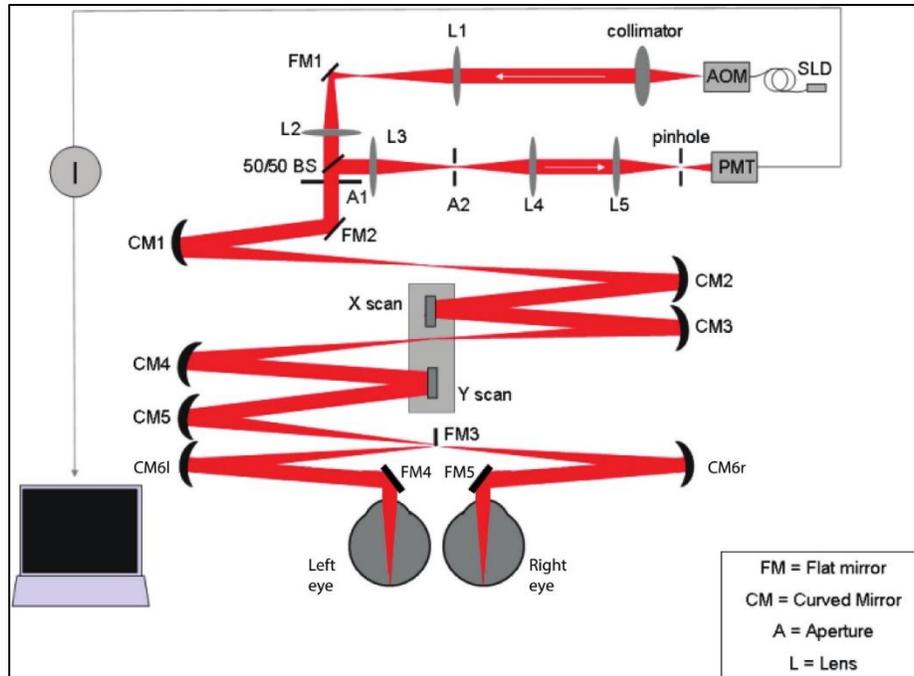


Figure 3.6. Schematic layout of the binocular modification to the Tracking Scanning Laser Ophthalmoscope. Mirror FM3 splits the field to left and right eye. The result is a split field image with left and right retinal images side by side in each frame. Eyes are not shown to scale (source: Stevenson et al., 2016).

3.2 Fundamental and derived metrics

In eye tracking research, fixations and saccades are the primary measurements used to analyze eye movements (Duchowski, 2007). Fixations are periods when the eyes remain still, focusing on a specific location or object. They indicate points of visual attention and provide insights into areas of interest or cognitive processing. Fixations are typically measured in terms of duration, meaning how long the eyes stay fixated on a specific location and location, meaning where the fixation occurs on a display or map. Saccades are rapid eye movements that occur between fixations. They involve quick shifts in gaze from one point of interest to another. Saccades help reposition the eyes to explore different areas of a visual stimulus and gather new visual information (Poole & Ball, 2005).

Metrics that are derived from fixations and saccades include gaze and scan path. Gaze refers to the overall pattern of eye movements over a given period, combining fixations and saccades. It represents cumulative visual attention and provides insights into the distribution and intensity of attention across a display or map (Poole & Ball, 2005). A scan path is the sequence of fixations and saccades made by an individual while observing a visual stimulus. It reveals the specific order and pattern in which different areas or objects are attended to. Analyzing scan paths can help identify individual strategies or preferences in visual exploration (Duchowski, 2007).

Additionally, pupil size and blink rate are measures studied by researchers to gain insights into visual attention and user behavior during tasks like map reading or interacting with online map services. Pupil size is the measurement of the diameter of the pupil. It can provide information about cognitive processes, arousal, and cognitive load. Changes in pupil size may indicate variations in attention, emotional response, or mental effort. Blink rate refers to the frequency at which individuals blink their eyes. It can be an indicator of visual fatigue, concentration, or attention levels. Changes in blink rate may reflect shifts in cognitive states or external stimuli.

3.3 Environmental conditions

Direct sunlight has a significant impact on the quality of data obtained from video-based P-CR eye trackers. Hansen & Pece (2005) and Holmqvist & Andersson (2017) provide examples demonstrating how infrared radiation from sunlight and hot light bulbs can undermine tracking in video-based P-CR systems. The importance of a controlled lighting environment is highlighted by (Y. Wang et al., 2010), who had to exclude 32% of participants recorded while driving a real car due to poor data quality, compared to only 17% of participants recorded in a car simulator. Even minor fluctuations in light levels can indirectly affect data quality. Several studies have identified the existence of the pupil-size artifact, where changes in pupil size impact the accuracy of gaze position in video-based P-CR systems (Choe et al., 2016; Hooge et al., 2019; Jaschinski, 2016; Wyatt, 2010).

Manipulating light levels to affect pupil size typically leads to increased gaze inaccuracy of 1 to 5 degrees. The reason behind the impact of pupil-size changes on reported gaze direction is that the pupil constricts and dilates asymmetrically, altering the shape of the pupil and causing a shift in the calculated center of the pupil image. In video-based P-CR eye trackers, this shift implies a change in gaze direction, even though the eyeball has not rotated relative to the head. The geometry of the setup, gaze direction, and distance to the eye camera have also been found to influence the magnitude of pupil-based errors (Wyatt, 2010). Furthermore, pupil size in P-CR eye trackers has been reported to be related to certain eye movement measures, such as saccadic peak velocity (Nyström et al., 2016). Accuracy in video-based P-CR trackers is generally better for participants with smaller baseline pupils measured under controlled illumination, as reported by Holmqvist (2017).

The pupil-size artifact may also affect other measures. For example, Hooge et al. (2019) found that light levels impact vergence estimations, with an error of 0.36-0.75 degrees per millimeter change in pupil size (similar findings were reported by Jaschinski, 2016). It can be expected that gaze position

errors caused by the pupil-size artifact will propagate to many other measures such as AOI (area of interest) analysis and other higher-order measures.

Vibration sources present in the recording environment contribute to increased variability in the gaze signal. In laboratory studies, nearby factors like an elevator shaft, a powerful air conditioning unit, or vibrations caused by people walking on hard floors in proximity can introduce measurable noise into sensitive eye tracking recordings. Sound in the recording environment is another source of oscillation that could cause the eye tracker to vibrate and affect the recorded data's quality. However, Hooge et al. (2019) conducted a study where they recorded Tobii TX300 data at an indoor science festival with moderately loud music and found accuracy values that closely matched the manufacturer's specifications. Despite these observations, controlled studies specifically examining the effect of vibrations on the quality of eye tracking data appear to be lacking.

The presence of other individuals during eye tracking recordings can have an impact on measures of eye movements and gaze behavior, but this effect is not yet fully understood. Social factors, such as social appropriateness, can influence participants' head and eye movements. When individuals are aware that they are being recorded by an eye tracker, they may adjust their gaze to focus only on socially appropriate stimuli (Nasiopoulos et al., 2015). Additionally, (Oliva et al., 2017) found longer latencies in the antisaccade task when participants were recorded in close proximity to one another, although the underlying reasons for this phenomenon remain unclear.

Chapter 4 : Experimental study design and implementation

4.1 Research design and approach

This empirical study is designed to explore how people visually perceive and pay attention to cartographic backgrounds used in online map services. It aims to investigate the effectiveness of cartographic symbolization elements and their impact on the process of map reading. By conducting an eye tracking experiment using the P-CR method and analyzing eye movements, the study seeks to identify key indices and locate the salient objects in the visual stimuli. Determining standard indices to describe visual perception on cartographic backgrounds, along with the assessment of overall observational behaviors, comparisons can be made between visual scenes. Insights into unique characteristics and challenges can be obtained concerning the map reading process. The contribution of this study in the field of cartography is the deeper understanding gained regarding the visual perception in web maps and the modeling of the visual behavior when observing such stimuli. The objective evaluation of these maps can examine their design, allowing its improvement to enhance user experience and usability. Modeling visual behavior contributes inputs to artificial intelligence systems, which predict the response of the observers, resulting to user-oriented maps.

4.2 Experimental Design

4.2.1 Description of the eye tracking system

The eye tracker used is the EyeLink[®] 1000 Plus eye tracker. A versatile device with its own programming interface and output data files (Eyelink Data File, EDF). The equipment used involves two PCs and two monitors. One PC is used as the display, having the Eyelink software installed along with the SR Research's tools and the other PC is used as the host where the host application starts and all the computations of true gaze and eye - motion events detection and analysis take place. The host monitor (Primary Monitor) is operated by the operator of the experiment and the display monitor (Secondary Monitor) serves the projection of the visual stimuli in its native resolution. Both PCs communicate via Ethernet. Using a timing – sensitive operating system, delays are minimized when acquiring and transmitting data. Real–time feedback during calibration or recording is also available.

Chapter 4 – Experimental study design and implementation

A high-speed camera is equipped, which in this experiment is placed under the display monitor on the desktop mount, without the need of a head stabilization, recording gaze data up to 2000 Hz. The host PC communicates with the Eyelink camera using a Gigabit network connection. Next to the camera, there is an infrared illuminator with its light emitting an LED wavelength under 940 nm (850 nm – 940 nm), combined with a filter to reduce or remove residual eye response. For the tracker to locate the subject even if the pupil image is lost, a target sticker is placed right above the eyebrows. The default setting of the camera lens for remote binocular tracking is 16 mm with a sample rate of 1000 Hz. In this study, a 25 mm lens is used with a sample rate of 500 Hz. The eye tracking system, with all its components, is presented in Figure 4.1.

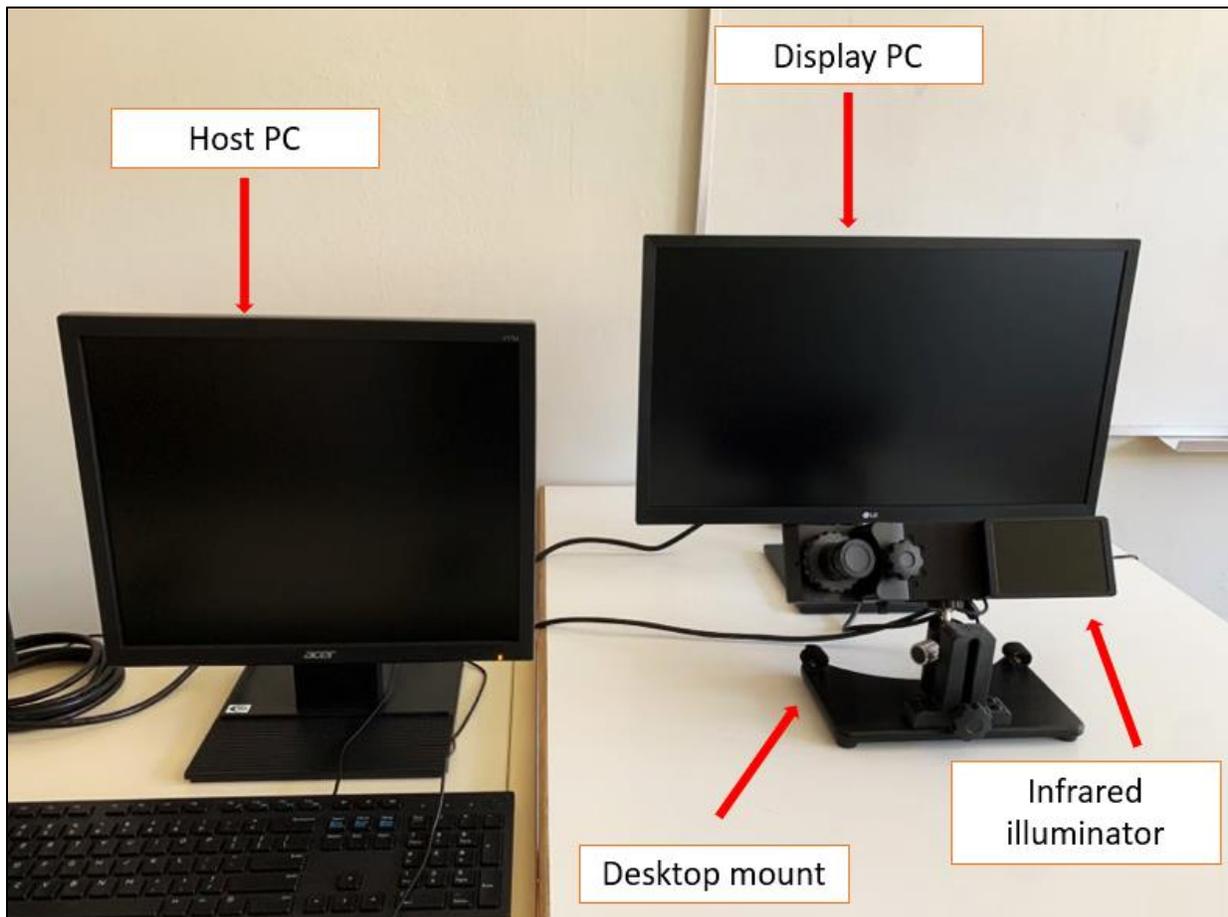


Figure 4.1. Eye tracking system (host PC on the left, display PC on the right and the desktop mount eye tracker beneath).

Additionally, this tracker can be combined with the SR Research Experiment Builder (SREB), a tool that integrates into the Eyelink software and hardware. The SREB utilizes visual programming,

Chapter 4 – Experimental study design and implementation

and it is quite flexible allowing many ways to create an eye tracking experiment as one judges in terms of the goal of one's study. Moreover, the EDF exported after a successful recording, can be inserted to the EyeLink® Data Viewer which is a tool for visualizing and filtering the collected data and creating output reports too.

The SR Research Experiment Builder (SREB) helps create the main structure of the conducting experiment. In every part that the workflow consists of, a pattern of action followed by a trigger and again a new action is used. This sequence allows the creation of a hierarchically organized flow diagram. The GUI contains all windows of the application allowing the user to control the parameters of the devices used, the parameters of the experiment's components categorized as 'action components' and 'trigger components and the parameters of the blocks and nodes inserted in the experiment's structure. The advantage of the visual programming is that dragging nodes of different properties and functions on the Graph Editor Window, nesting different sequence blocks, and connecting them, generates the experimental flow diagram. In Figure 4.2, the GUI is depicted showing the components of a demo experiment linked to each other to form its structure.

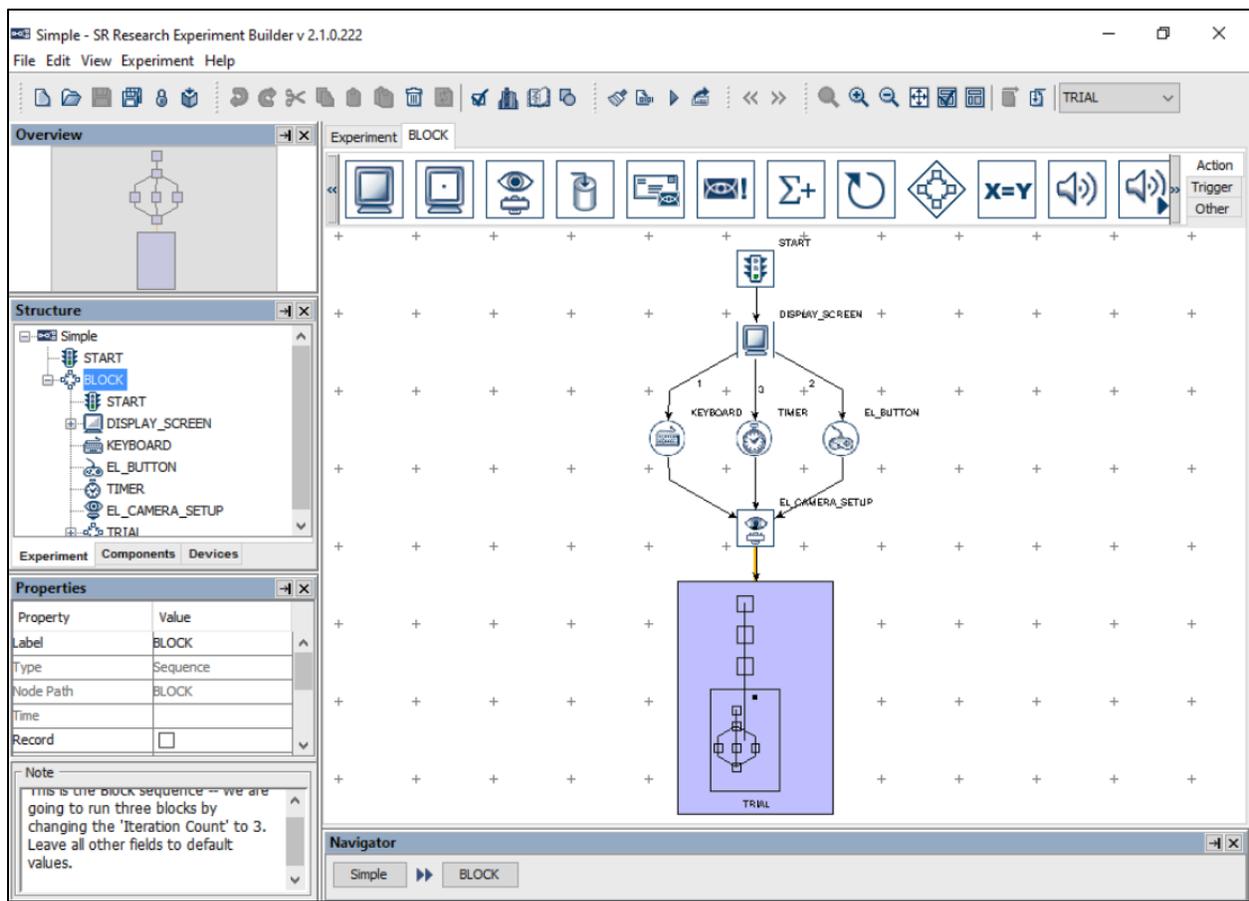


Figure 4.2. SR Experiment Builder user interface (source: SR Research Ltd, 2020)

Chapter 4 – Experimental study design and implementation

The EyeLink® Data Viewer is the tool that allows data manipulation derived from the conduction of the experiment. The user can visualize the fixations and the saccades on heatmaps measuring the number of fixations and the maximum duration of the fixation in a trial and export various reports with information and events about these metrics for every trial (SR Research Ltd, 2021b). Moreover, the data can be filtered by applying thresholds depending on the preferences and the tolerance needed for the project. Aggregate data for all trials in an experiment and aggregate data reports are available. In addition, grouping trials from different participants, the visualization of aggregate data on a particular image is possible, as well as summary statistics for an entire trial group, tools that help the upcoming analysis. The EyeLink® Data Viewer user interface is presented in Figure 4.3.

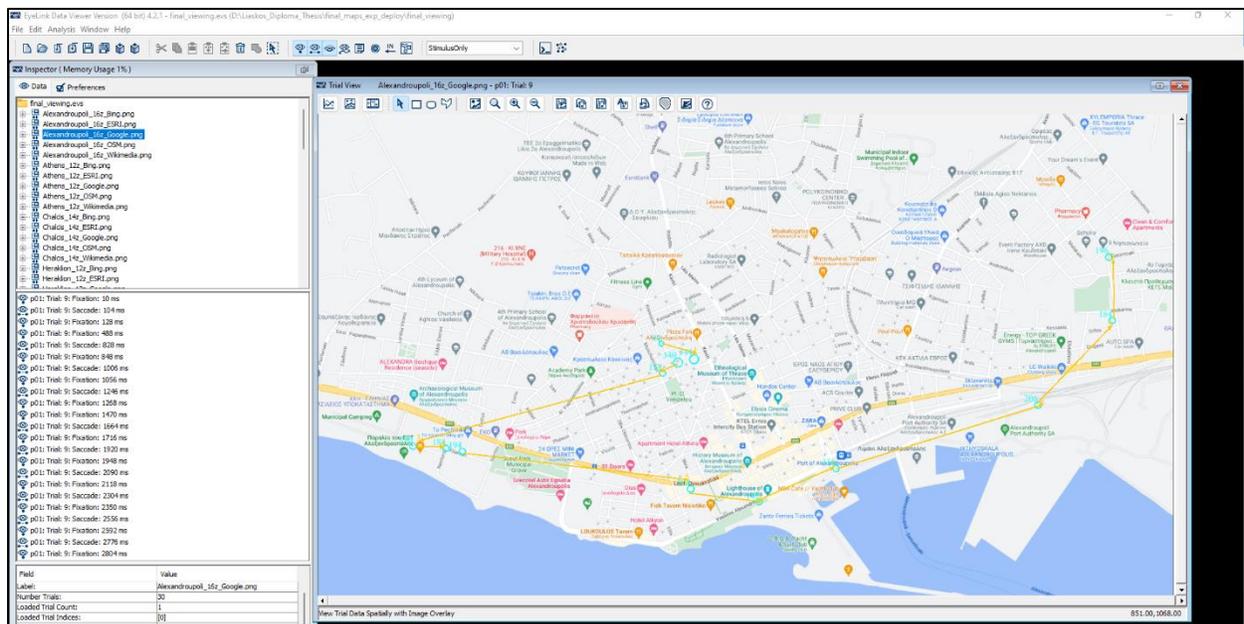


Figure 4.3. Data Viewer user interface.

4.2.3 How the eye tracking system works

The EyeLink® 1000 Plus camera is attached to a Desktop Mount, making this combination the eye tracking system used in this experiment. This system, in binocular remote recording, employs the pupil – center (P – CR)/corneal – reflection method to measure the point of regard during eye movement analysis. An infrared light emitted by an LED embedded in the infrared camera is directed into the eye. This light creates strong reflections in certain target eye features, making them easier to track without causing discomfort or visual distraction to the user. The infrared light enters the retina, and a sizable portion of it is reflected, resulting in the pupil appearing as a bright and well-defined disc. The corneal reflection, on the other hand, is generated by the infrared light and appears as a small

Chapter 4 – Experimental study design and implementation

glint (Goldberg & Wichansky, 2003). By measuring the vector between these two points and employing trigonometric calculations, the point of regard can be determined. This method enables the tracking of eye movements independently of head movements. By using the corneal-reflection/pupil-center method in eye tracking, researchers can gain insights into where individuals are looking on a display or map, providing valuable information about visual attention, information processing, and cognitive processes during tasks such as map reading or interacting with online map services (Duchowski, 2007).

The eye tracker applies the Ellipse Fitting algorithm to track the pupil. This mode calculates the center of the pupil by fitting an ellipse based on the thresholded pupil mass. On the occlusion of the pupil, the algorithm interpolates points that may drop behind the eyelid or the eyelashes and, in this way, achieves a better approach on the pupil position. This mode also decreases drift potential and copes with pupil occlusion, giving a more accurate estimation, but there is a higher noise level compared to the Centroid mode which is used on other tracking methods (SR Research Ltd, 2021a).

Having set up the tracker to detect pupil and pupil corneal position, the next crucial step is the association between the pupil position in the camera image and the gaze on the display screen. Calibration is the process in which the coordinate system of the observer is associated with the coordinate system of the monitor displaying the stimuli. The participant fixates on targets positioned by the system on known locations. For every target recorded, the positions of the pupil corneal and the target are used to compute gaze positions. The default system calibration uses nine targets but there is an option of thirteen points for larger areas and for remote tracking methods. Every target is changing position one second after a stable fixation when the “Auto Trigger” option is enabled. The targets are shown in random order. If a grid is formatted, the calibration is successful, otherwise the calibration is repeated. Calibration is performed for every participant, and the demanded accuracy is determined by the experiment type (SR Research Ltd, 2021a).

Immediately after calibration, the system can predict gaze position from the pupil position and the accuracy can be tested by running a validation. Targets are shown to the observer, and they change location after a stable fixation. Calibration accuracy, in this procedure, is tested on uncalibrated points too. Errors gathered from data in calibration procedure, derive from the eye tracking system and eye movement controls, but given that the eye tracking system has extremely low pupil-position noise and remarkably high resolution, these errors are eliminated. A fact which shows that inaccuracies can be recorded if there is a small decline while the observer is fixating on the targets or is not sitting still during calibration.

4.2.4 Specifications

Table 4.1. Eye tracker specifications (source: Eyelink 1000 plus user manual).

Desktop and LCD Arm Mounts	
Remote Tracking	
Average Accuracy	0.25 – 0.5° typical
Sampling Rate	Binocular: 250, 500, 1000 Hz
End – to – end Sample Delay	500 Hz: M = 3.29 ms, SD = 0.58 ms
Blink / Occlusion Recovery	2.0 ms @ 500 Hz
Spatial Resolution	0.01°
Noise with Participants	Filter (Off/Normal/High) 500 Hz: 0.03°/0.02°/0.01° (25 mm lens)
Eye Tracking Principle	Dark Pupil – Corneal Reflection
Pupil Detection Model	Ellipse Fitting
Pupil Size Resolution	0.1% of diameter (25 mm lens)
Gaze Tracking Range	Customizable Default is 32° horizontally x 25° vertically
Allowed Head Movements Without Accuracy Reduction	25 mm lens: 22 x 22 cm at 60 cm 25 x 25 cm at 70 cm
Optimal Camera – Eye Distance	40 – 70 cm
Infrared Wavelength	850 – 940 nm
Glasses Compatibility	Good

4.2.5 Experimental Stimuli

The visual stimuli used in the experiment are static maps, taken as excerpts from the web maps provided from five different online map services (Google Maps, Wikimedia, Bing maps, ESRI, OSM). The excerpts need to be in the native resolution of the screen (1920 x 1080 pixels). The maps depict various cities of Greece, in different zoom levels, based on their population. The higher the population, the lower the zoom level. Population data are provided by the Hellenic Statistical

Chapter 4 – Experimental study design and implementation

Authority (ELSTAT). It is decided that there will be 3 different zoom levels, and on every zoom level there will be five cities. Table 3 contains the cities for which cartographic backgrounds were exported.

Table 4.2. Cities depicted in the cartographic backgrounds.

12z	14z	16z
Athens	Chalcis	Alexandroupoli
Heraklion	Ioannina	Kalamata
Larissa	Serres	Katerini
Patras	Trikala	Kavala
Thessaloniki	Volos	Xanthi

Using QGIS software, 75 maps are exported. All web map services use the same coordinate reference system, the WGS84 ellipsoid, and the “Web Mercator” projection, introduced by Google Maps, which is a new version of the Mercator Projection (Skopeliti & Stamou, 2019). Using Google Maps, two points diagonally from each other are selected. The position of the points can be seen in Figure 4.5. Their coordinates are extracted and they serve as boundaries of the cartographic backgrounds. The coordinates of the points are inserted on .txt files. For every city that belongs in a zoom level category, a set of points is inserted into the GIS application. Using the Quick Map Services plug in, base maps from different online map services are available on the extent set from the points. As presented in Figure 4.4, choosing the extent (scale) and the resolution value, along with the output height and width, the cartographic backgrounds are exported for every service. The height and width should match the dimensions of the display screen.



Figure 4.4. Output image parameters regarding the resolution, height and width.

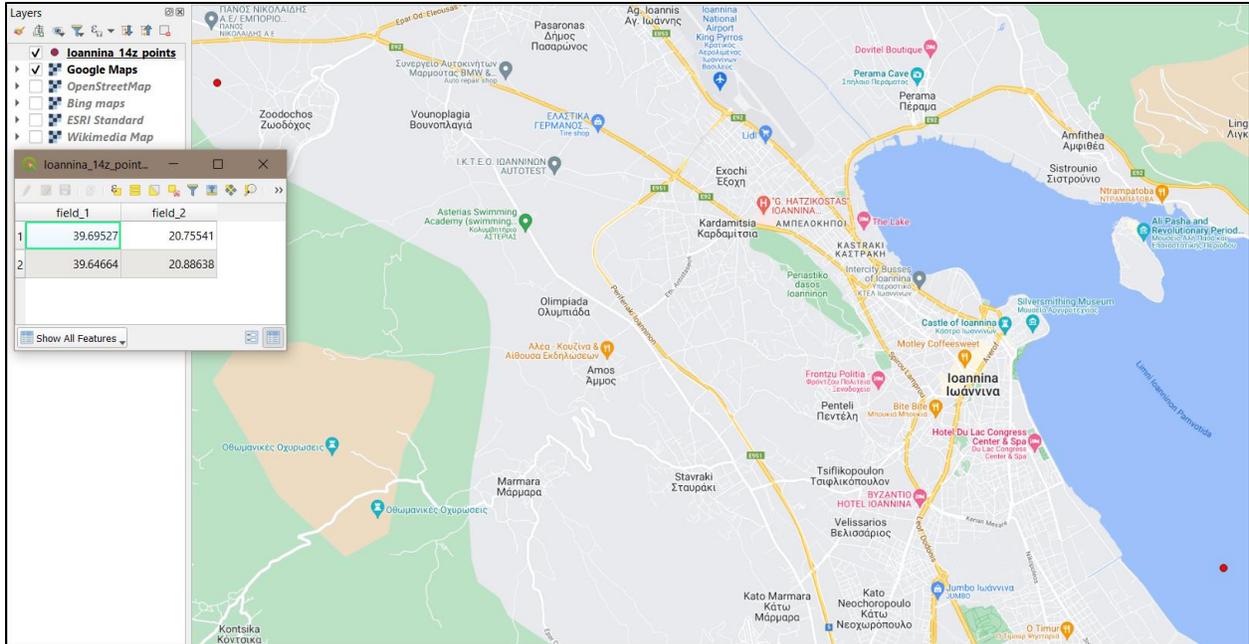


Figure 4.5. Cartographic backgrounds in the QGIS environment. The red points are the boundaries for the export image and the attribute table contains their coordinates in WGS 84'.

4.2.6 Participants

Thirty participants involved in the experiment, with 53% being female and 47% being male. 73% has right dominant eye and 27% has left dominant eye. The ages of the participants varied with a range from 19 to 61 years old, with most participants being ages 21 to 23 years old. None of them had any experience with eye tracking experiments. The gender of the participants is depicted in Figure 4.6, the participant's dominant eye is presented in Figure 4.7, and the age of the participants appears in Figure 4.8.

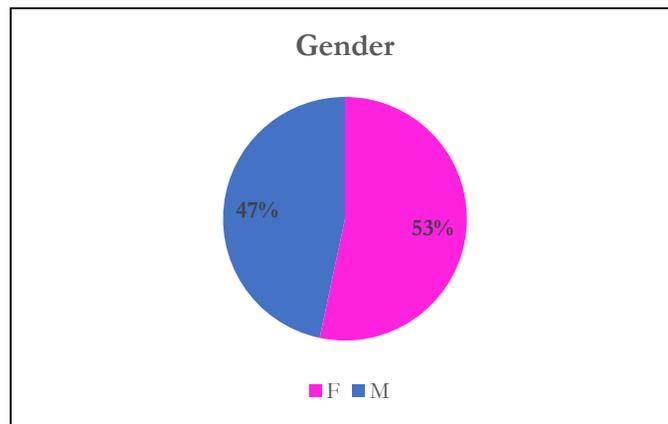


Figure 4.6. Gender of the participants.

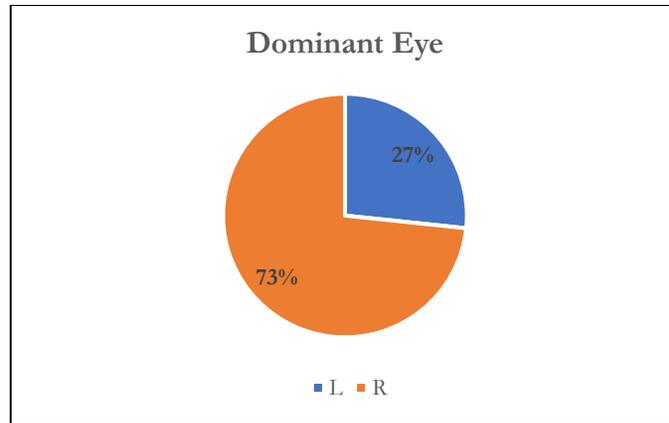


Figure 4.7. Participants' dominant eye.

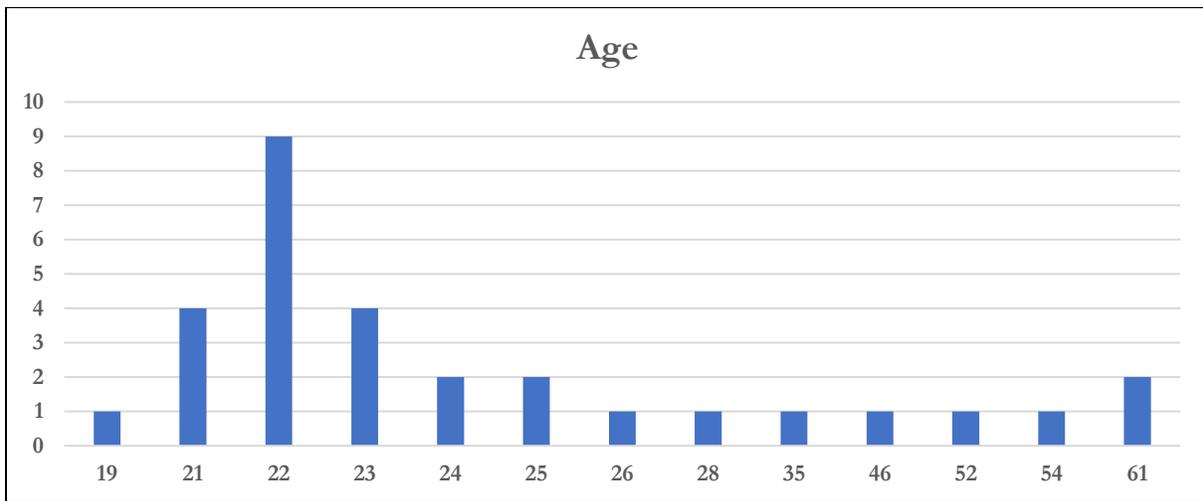


Figure 4.8. Participants' age.

4.3 Experimental Programming

The structure of the experiment is designed with the SR Research Experiment Builder. The main experiment flow diagram consists of three main blocks nested in one another, following the hierarchical conditions: the Sequence Block, the Trial Block, and the Recording Block. The first node added is the Camera Setup, which enables the procedure of calibration, validation and all the steps to set the Desktop Mount's camera according to every participant at the start of the experiment. The Sequence Block contains the Drift Correction node, the Trial Block contains the Prepare Sequence node and the data source of the experiment, and the Recording Block contains the Display Screen, followed by a Timer and a blank Display Screen with its own timer. Figure 4.9 shows the structure of the experiment in the Experiment Builder's user interface.

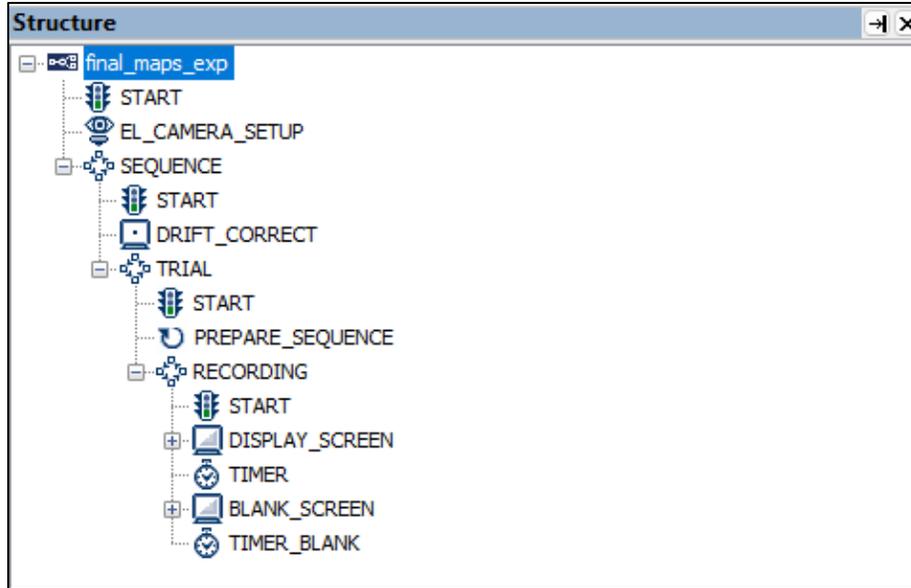


Figure 4.9. Structure of the main experiment in SR Experiment Builder UI.

Considering that calibration accuracy must be maintained throughout the experiment, visual stimuli should be presented so drift checks can be made. Therefore, the trial data source is divided into three parts, 25 maps each. Drift correction happens in between these parts, and if the accuracy is not satisfying, calibration can be performed again. For that reason, the Sequence Block iterates three times with every iteration representing a part of the data source. The Trial Block contains the data to be recorded in real time, so recording is enabled, and the Sequence Block serves to draw the images from the source to the display screen.

The images representing the visual stimuli are stored on the display pc, the experiment graph file, and its folder, and they must be connected to the experiment. The data source in the Trial Block is a list containing the id and the name of the images. To display the images on the display screen, the files stored in the pc are linked with their directories through the data source based on their name. When clicking on the builder’s interface on the Display Screen, you pick the position of the image on the screen and the display order which is randomized to mitigate order bias and produce fair results. Every image is displayed for three seconds and is followed by a blank screen with a neutral background lasting for one second, to make the observer start over the exploration from the starting point, the center of the image. The time of the stimuli is determined by the total duration of the experiment in order to avoid the physical fatigue that can occur to the participant. The decision for the specific duration of each stimulus, is the result of two trials with longer durations. The initial duration for every stimulus was five seconds, resulting to a total duration of seven and a half minutes of visual exploration, excluding the camera setup. The second trial, proceeded with each stimulus lasting four

seconds, making the total duration equal to six minutes. Eventually, the duration is set to three seconds, making the total duration equal to five minutes.

4.4 Recording phase

4.4.1 Environmental conditions

When recording the samples of the eye movements, it is important to have an efficient experiment design, but also to consider the external factors that have an effect while recording the data. The experiment takes place in a controlled environment, with light conditions, environmental vibrations, ambient noise and presence of others manipulated in a manner to achieve the best possible measurements of the gaze data and the various metrics.

The setup of the eye tracking system is located in a room where the environmental factors can be controlled. Individuals are not allowed inside during the recording, except for the participant and the experiment manager. There is moderate ambient natural lighting, and in some cases with participants wearing glasses or mascara the lighting is moderated more. The experiment is conducted in silence and the participant is focused on the visual stimuli. All the questions are answered before the experiment, when the participant is signing the consent form that states all the information about the experiment.

4.4.2 Experimental procedure

The experiment needs the approval of the Research and Ethics Committee of UNIWA, from which a consent form, consisting of four pages, is given for every volunteer to read and fill, in order to participate in the experiment. The consent form describes the objective of the empirical study, along with the experimental process and how the data acquired will be used. Afterwards, if the volunteer has agreed to participate, a form is filled with the name, age, gender, experience on eye tracking experiments and after the completion of the experiment, the dominant eye.

Before starting the experiment, the participant is getting accustomed to the process via a demo experiment. The demo experiment is important because it familiarizes the participant with the process of his position setup, calibration, validation, and the form of the visual stimuli to be observed. Moreover, the calibration process can be conducted more than once to achieve the best possible accuracy. The participant must be in a comfortable position as it will not change during the

Chapter 4 – Experimental study design and implementation

experiment, and it should not, because the accuracy will decrease, and calibration should occur again. In Figure 4.10, the visual stimuli used in the demo experiment are presented.

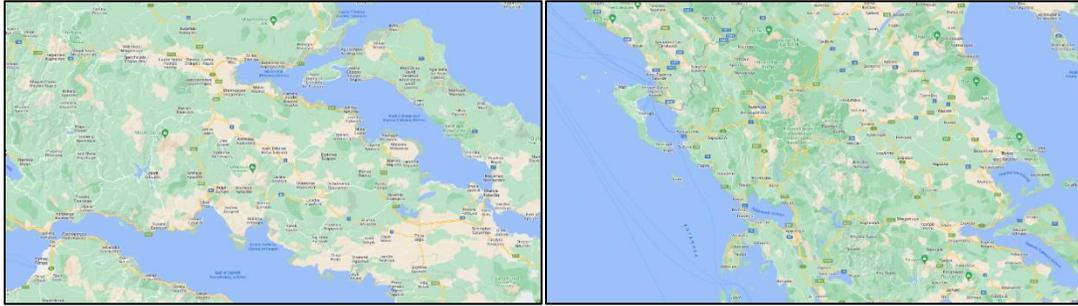


Figure 4.10. Visual stimuli used in demo experiment. Cartographic background excerpts from Google Maps.

Designing the experiment, one should consider the external variables affecting it. The participants need to be focused on the task. Furthermore, there are specifications to be followed depending on the equipment that is used. To accomplish the optimal average accuracy given by the manufacturer (table 1), typically 0.25° – 0.5° of the visual angle, the participant's position must always be stable. The display monitor and eye tracking device should not be moved, and the observer is to be placed 50 cm to 60 cm from the desktop mount. Once the participant is set to about 1.75 times the display width and the whole monitor is visible without surpassing 32° of visual angle horizontally and 25° vertically, the experiment can begin. If the participant is not stabilized, the data is to be discarded. Figure 4.11 visualizes the geometry that involves the participant and the eye tracking system.

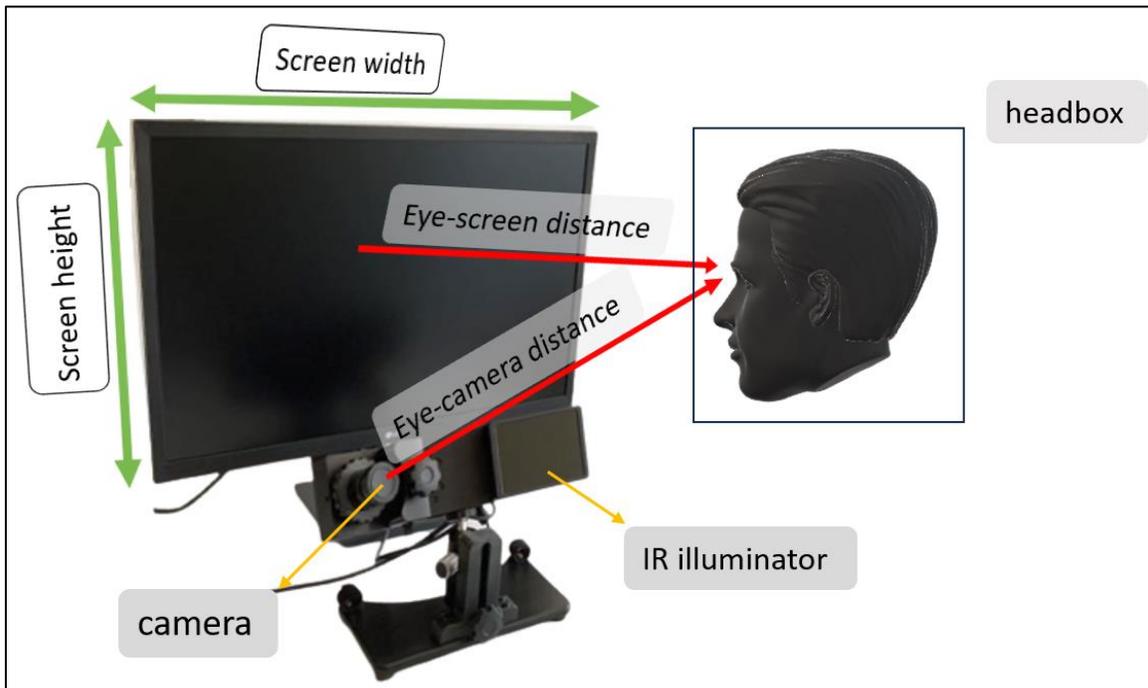


Figure 4.11. Eye tracking system geometry.

Chapter 4 – Experimental study design and implementation

Once the preparations are made, calibration and validation, are required. The targets for the calibration are of known location and have a specific color, size and shape. The display screen coordinates are listed in the SR Research Experiment Builder (SREB) Manual. The size of the target is defined by two parameters: the outer size and the inner size. The default target is a filled circle with a central hole, with the outer size specifying the diameter of the outer circle and the inner size specifying the diameter of the inner circle. The inner circle aids the accuracy of the calibration. The size selected for the outer circle for peripheral detectability is set to 17 pixels and the size of the inner circle is set to 5 pixels. The chosen color filling is red (RGB 255, 0, 0) and the inner circle color is white (RGB 255, 255, 255). The background color is of a grey shade (RGB 230, 230, 230), so the contrast created is adequate. In Figure 4.12, the location of the drift correction points is shown, along with the color chosen for the targets depicted in Figure 4.13, which constitute the calibration targets.

X Location	960
Y Location	540
Apply Correction	CURRENT
Allow Setup	<input checked="" type="checkbox"/>
Draw Drift Correction Target	<input checked="" type="checkbox"/>
Clear Target At Exit	<input checked="" type="checkbox"/>
Foreground Color	
Background Color	

Figure 4.12. Location of the drift correction target and the chosen color for the targets of calibration, validation and drift correction (RGB [255 0 0]).

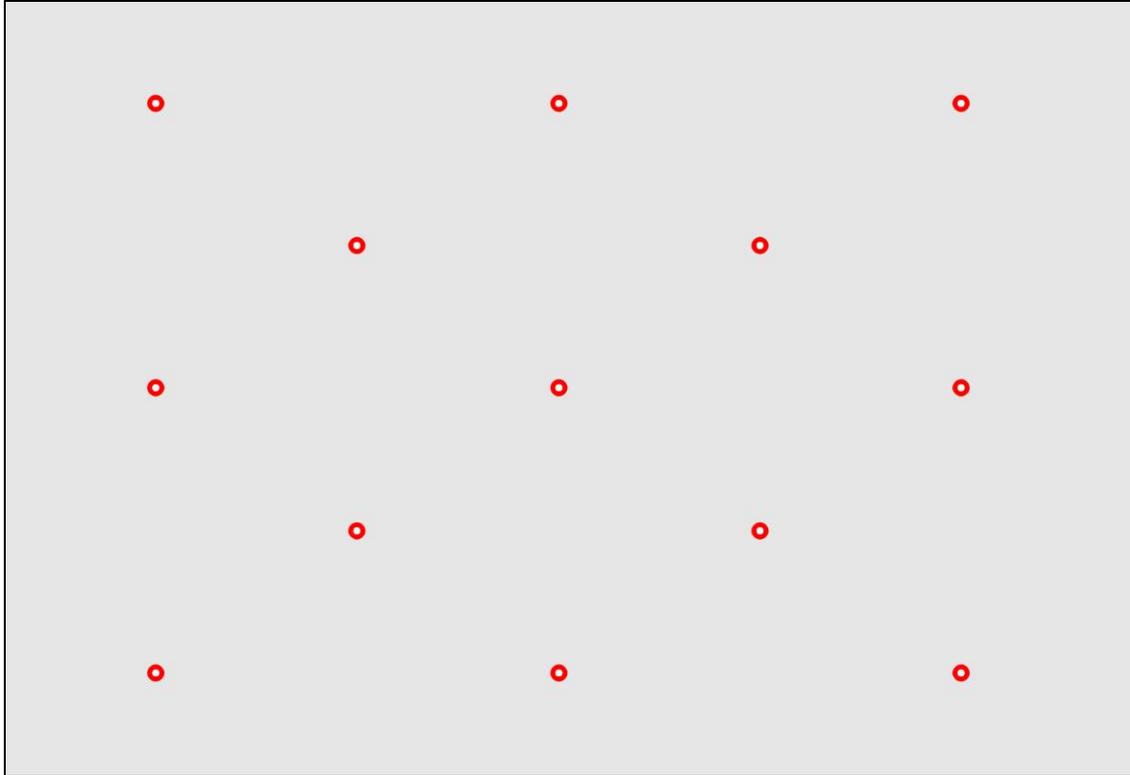


Figure 4.13. Calibration targets: position, size, color and background color.

The main experiment total duration is calculated to be around six minutes and the demo experiment's duration to be around one minute, both calculated without having to set up the participant's position and in ideal conditions. Calibration and validation require a minute, there are three drift corrections in between the three divisions of the data source, and there is a blank screen after ever stimulus appearance. Conducting the experiment in free viewing conditions, there are no questions to be answered about the stimuli, in order to achieve the objective tracking of the participants' eye movements, without them being premeditated.

4.5 Data Analysis

4.5.1 Participants' data

Every participant's session is saved on an EyeLink data file format (EDF). EDF contain the recorded eye movements, messages sent during the experiment regarding the beginning and the completion of the camera setup and every trial recording. EDF files can be opened in EyeLink® Data Viewer to extract the metrics needed for the analysis. The file created by opening every participant's EDF file in the Data Viewer is saved as a viewing session. Consequently, there can be a viewing session

Chapter 4 – Experimental study design and implementation

with all participants' data to manipulate them in a cumulative manner and produce aggregated visualizations of the gaze data. In Figure 4.14, the fixation and saccade events are depicted in the Data Viewer user interface. In Figure 4.15, fixations, saccades, and blink saccades, are all presented on top of a visual stimulus in aggregate mode, meaning that the events visualized are cumulative.

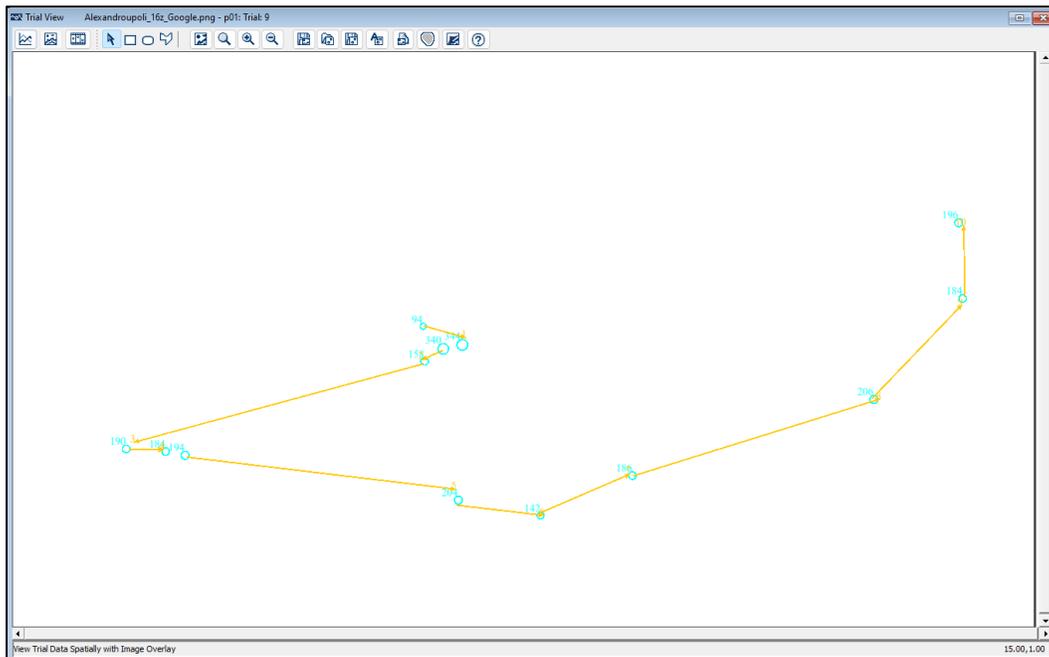


Figure 4.14. Fixations (cyan) and saccades (yellow) in the EyeLink® Data Viewer UI.

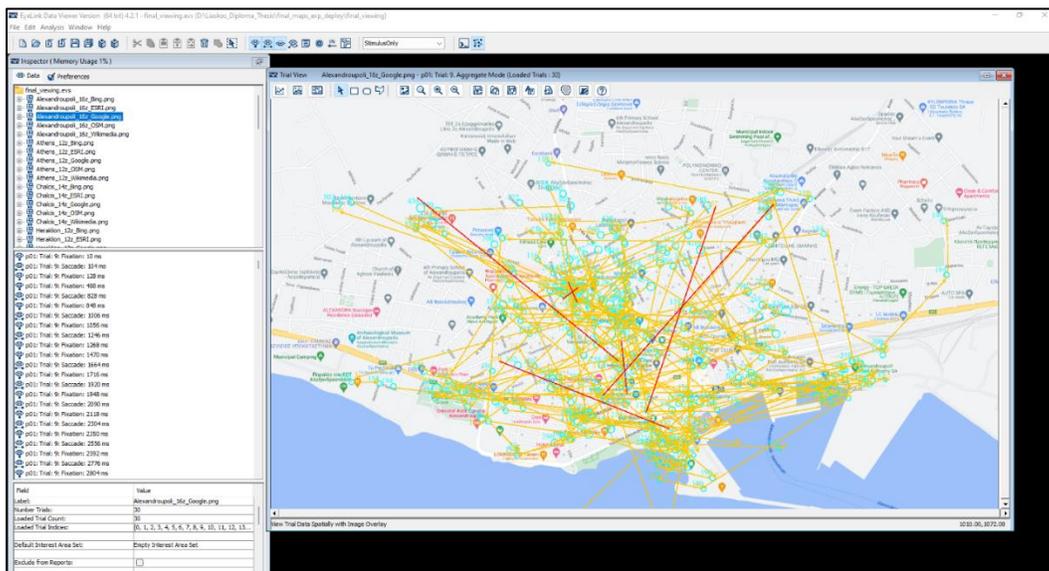


Figure 4.15. Fixations (cyan), saccades (yellow), blink saccades (red) overlaid on a random visual stimulus in the EyeLink® Data Viewer UI.

Chapter 4 – Experimental study design and implementation

Having the data organized in the Data Viewer interface, the next step is to export three different reports for the analysis. The first report contains the raw gaze data organized in ten columns. As it can be seen in Figure 4.16, the first column is the recording session label, referring to the participant, the second on is the trial index and is quite helpful to understand the real order of every stimulus, because the order of appearance is randomized during the experiment. The third column is the stimulus name, the fourth and fifth columns are the average gaze coordinates, the sixth and seventh are left gaze coordinates and eighth and ninth are the right gaze coordinates. The second report contains the fixations report with the first three column being the same as the raw gaze data report, and the next two being the total number of fixations for each trial and the duration of every fixation in all fixations. The fixations report is presented in Figure 4.17. The third report refers to the saccades report, containing the same three first columns, and the fourth column is the saccade amplitude of every saccade. In Figure 4.18, the saccades report is shown.

	RECORDING_SESSION_LABEL	TRIAL_INDEX	TIMESTAMP	image	AVERAGE_GAZE_X	AVERAGE_GAZE_Y	RIGHT_GAZE_X	LEFT_GAZE_X	RIGHT_GAZE_Y	LEFT_GAZE_Y
1	p01	1	1500258.00	Volos_14z_OSM.png	947.30	573.50	957.20	937.40	580.00	567.00
2	p01	1	1500260.00	Volos_14z_OSM.png	947.35	573.65	957.30	937.40	579.90	567.40
3	p01	1	1500262.00	Volos_14z_OSM.png	947.45	573.85	957.50	937.40	579.90	567.80
4	p01	1	1500264.00	Volos_14z_OSM.png	947.55	574.10	957.70	937.40	580.00	568.20
5	p01	1	1500266.00	Volos_14z_OSM.png	947.65	574.25	957.90	937.40	580.10	568.40
6	p01	1	1500268.00	Volos_14z_OSM.png	947.75	574.20	958.10	937.40	580.10	568.30
7	p01	1	1500270.00	Volos_14z_OSM.png	947.90	574.05	958.40	937.40	580.00	568.10
8	p01	1	1500272.00	Volos_14z_OSM.png	947.90	574.00	958.50	937.30	580.00	568.00
9	p01	1	1500274.00	Volos_14z_OSM.png	947.90	573.95	958.60	937.20	580.10	567.80
10	p01	1	1500276.00	Volos_14z_OSM.png	947.75	573.95	958.40	937.10	580.20	567.70
11	p01	1	1500278.00	Volos_14z_OSM.png	947.65	573.90	958.30	937.00	580.30	567.50
12	p01	1	1500280.00	Volos_14z_OSM.png	947.55	573.95	958.10	937.00	580.60	567.30
13	p01	1	1500282.00	Volos_14z_OSM.png	947.50	574.05	957.90	937.10	580.90	567.20
14	p01	1	1500284.00	Volos_14z_OSM.png	947.45	574.15	957.70	937.20	581.20	567.10
15	p01	1	1500286.00	Volos_14z_OSM.png	947.50	574.20	957.60	937.40	581.40	567.00
16	p01	1	1500288.00	Volos_14z_OSM.png	947.50	574.30	957.40	937.60	581.60	567.00
17	p01	1	1500290.00	Volos_14z_OSM.png	947.55	574.35	957.40	937.70	581.70	567.00
18	p01	1	1500292.00	Volos_14z_OSM.png	947.55	574.50	957.30	937.80	581.90	567.10
19	p01	1	1500294.00	Volos_14z_OSM.png	947.55	574.60	957.20	937.90	581.90	567.30
20	p01	1	1500294.00	Volos_14z_OSM.png	947.55	574.60	957.20	937.90	581.90	567.30

Figure 4.16. Excerpt from the gaze data report exported from EyeLink® Data Viewer in a .txt file.

Chapter 4 – Experimental study design and implementation

	RECORDING_SESSION_ LABEL	TRIAL_INDEX	image	CURRENT_SAC_ AMPLITUDE
1				
2	p01	1	Volos_14z_OSM.png	4.16
3	p01	1	Volos_14z_OSM.png	5.49
4	p01	1	Volos_14z_OSM.png	1.5
5	p01	1	Volos_14z_OSM.png	3.19
6	p01	1	Volos_14z_OSM.png	1
7	p01	1	Volos_14z_OSM.png	9.95
8	p01	2	Kalamata_16z_Google.png	1.84
9	p01	2	Kalamata_16z_Google.png	3.37
10	p01	2	Kalamata_16z_Google.png	16.58
11	p01	2	Kalamata_16z_Google.png	5.41
12	p01	2	Kalamata_16z_Google.png	1.25
13	p01	2	Kalamata_16z_Google.png	23.69
14	p01	2	Kalamata_16z_Google.png	5.62
15	p01	2	Kalamata_16z_Google.png	1.08
16	p01	2	Kalamata_16z_Google.png	1.8
17	p01	2	Kalamata_16z_Google.png	20.73
18	p01	3	Kavala_16z_Wikimedia.png	2.05
19	p01	3	Kavala_16z_Wikimedia.png	10.79
20	p01	3	Kavala_16z_Wikimedia.png	2.87

Figure 4.17. Excerpt from the saccade report exported from EyeLink® Data Viewer.

	RECORDING_SESSION_ LABEL	TRIAL_INDEX	image	TRIAL_FIXATION_ TOTAL	CURRENT_FIX_ DURATION
1					
2	p01	1	Volos_14z_OSM.png	11	388
3	p01	1	Volos_14z_OSM.png	11	208
4	p01	1	Volos_14z_OSM.png	11	108
5	p01	1	Volos_14z_OSM.png	11	100
6	p01	1	Volos_14z_OSM.png	11	184
7	p01	1	Volos_14z_OSM.png	11	250
8	p01	1	Volos_14z_OSM.png	11	486
9	p01	1	Volos_14z_OSM.png	11	204
10	p01	1	Volos_14z_OSM.png	11	146
11	p01	1	Volos_14z_OSM.png	11	370
12	p01	1	Volos_14z_OSM.png	11	310
13	p01	2	Kalamata_16z_Google.png	13	252
14	p01	2	Kalamata_16z_Google.png	13	112
15	p01	2	Kalamata_16z_Google.png	13	130
16	p01	2	Kalamata_16z_Google.png	13	352
17	p01	2	Kalamata_16z_Google.png	13	160
18	p01	2	Kalamata_16z_Google.png	13	226
19	p01	2	Kalamata_16z_Google.png	13	184
20	p01	2	Kalamata_16z_Google.png	13	132

Figure 4.18. Excerpt from the fixation report exported from EyeLink® Data Viewer.

Chapter 4 – Experimental study design and implementation

Deciding which data are useful and which are to be discarded, is a process to be aware in the beginning of the data recording. Data collection started with achieving the best possible accuracy from each participant, translated to average accuracy according to the specifications, and max error less than a degree of visual angle, and even better, less than 0.8° . Furthermore, the in-between drift corrections maintained the accuracy throughout the whole process. The equipment made this task possible.

Before exporting the reports, the data must be filtered by applying thresholds on fixation and saccade metrics. Fixations last for 218 milliseconds on average, with a range of 66 to 416 milliseconds. The minimum time for a fixation is significant with researchers being advised to set the lower threshold to at least 100 ms, according to Inhoff & Radach (1998). The threshold is set to 80 milliseconds. The minimum saccade amplitude refers to the smallest distance that an eye movement must cover to be considered a saccade. It represents the threshold below which eye movements are classified as fixations rather than saccades. Commonly used thresholds range from 0.1 to 1 degree of visual angle. The threshold set is 1 degree of the visual angle. In Figure 4.19, the preferences tab is depicted, and more specifically, the data filters, in which the threshold values are set.

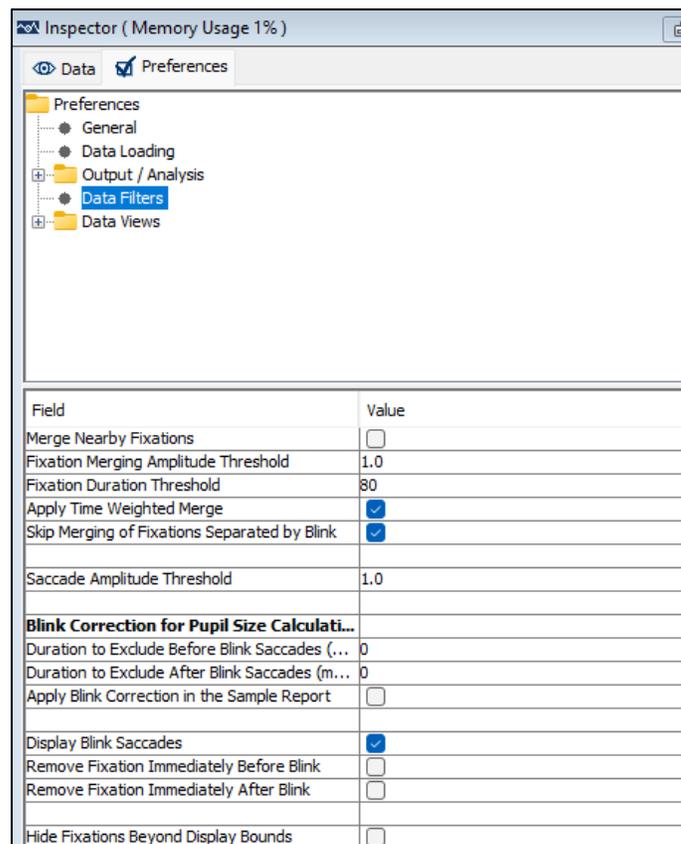


Figure 4.19. Data filters in the EyeLink® Data Viewer UI.

4.5.2 Fixation and saccade indices

The fixations and saccades reports can be further analyzed and provide insights on the manner the participants have observed the visual stimuli. For three different categories: 1) fixation duration, 2) total number of fixations, and 3) saccade amplitude, five indices are calculated: a) min, b) max, c) average, d) median, e) standard deviation (std). The minimum and maximum indices, for all three categories, show the range of the values, the average indice calculates the mean value, the median indice calculates the mean value excluding the outliers and the standard deviation indice is a measure of the dispersion of the sample values respecting to the mean value. The indices are calculated for all results, based on the stimulus, so later on, filters can be applied to extract information and make comparisons between different zoom levels, different services and even different cities.

The size of the data makes it difficult to manipulate, and for that reason MATLAB software (MathWorks®) is used to filter the reports and conduct the analysis. The indices are measured cumulatively for every image/stimulus. The reports export for every participant, every participant's fixation duration in every image/stimulus and every saccade's amplitude in every image/stimulus. Gathering all measures (fixation duration, total number of fixations and saccade amplitude) from all participants for every image, permits the creation of a new table, from which the indices are calculated for every category and they are exported in a .txt file as a string array. The string array appears in figure 4.20.

```

1 "Alexandroupoli_16z_Bing.png,82,2990,298.8444,229,259.6716"
2 "Alexandroupoli_16z_ESRI.png,82,1820,309.2374,250,217.0272"
3 "Alexandroupoli_16z_Google.png,82,2990,263.612,220,209.1946"
4 "Alexandroupoli_16z_OSM.png,80,1494,289.1136,228,201.6051"
5 "Alexandroupoli_16z_Wikimedia.png,84,2408,292.2418,240,215.1611"
6 "Athens_12z_Bing.png,82,1534,286.1367,229,185"
7 "Athens_12z_ESRI.png,82,2464,314.7812,255,252.2051"
8 "Athens_12z_Google.png,80,1476,265.7852,227,155.2919"
9 "Athens_12z_OSM.png,82,1982,308.5802,250,238.3135"
10 "Athens_12z_Wikimedia.png,84,2990,302.0604,238,283.0786"
11 "Chalcis_14z_Bing.png,82,1518,295.2945,248,197.0953"
12 "Chalcis_14z_ESRI.png,80,2066,311.2664,244,237.7655"
13 "Chalcis_14z_Google.png,80,2990,312.5447,234,305.7302"
14 "Chalcis_14z_OSM.png,82,2072,294.0441,237,215.2796"
15 "Chalcis_14z_Wikimedia.png,86,2990,300.0597,238,249.4371"
16 "Heraklion_12z_Bing.png,90,2460,295.1618,249,208.4432"
17 "Heraklion_12z_ESRI.png,80,2990,317.7874,243,280.3608"
18 "Heraklion_12z_Google.png,80,1668,294.684,226,220.2887"
19 "Heraklion_12z_OSM.png,86,2306,314.0863,252,238.7388"
20 "Heraklion_12z_Wikimedia.png,80,2990,318.2677,235,311.85"

```

Figure 4.20. Excerpt from the string array of the calculated indices.

Chapter 4 – Experimental study design and implementation

It is important to calculate those indices, as they facilitate the analysis of the variables of interest. Comparing the relations between the variables, can lead to meaningful conclusions. The minimum value is the smallest observation on the three categories selected, useful to identify outliers and understand the lower bound of the range. The maximum value represents the largest value, identifying the outliers and the upper bound of the range. The average value is a measure of central tendency, representing the typical value. The median value is the middle value in a sorted dataset, aiding as a measure of central tendency robust to outliers. The standard deviation value measures the variability, quantifying how spread out the values are around the mean value, with a higher standard deviation indicating greater variability.

The variables of interest are the eye tracking indices, obtained from the reports. The objective of their analysis is the understanding on how individuals allocate attention, how they explore the visual stimuli and the salient locations of the visual stimuli presented. Moreover, differences can be revealed among the participants, during map reading or maybe there is uniformity on the pattern of the processing strategies employed. Visualizing the indices in an effective way, such as bar charts, their interpretation can be facilitated and effectively communicate key insights and patterns.

The fixation duration and the total number of fixations refer to the fixations, and the saccade amplitude refers to the saccades. Larger number of fixations concentrated in an area, suggest that the objects of this area/visual stimulus attract the observer's visual attention, making them salient objects (Jacob & Karn, 2003); (Poole et al., 2005). The fixation duration indicates the difficulty of extracting information from a certain visual stimulus, with higher fixation duration meaning that the difficulty is increased (Jacob & Karn, 2003) or that the objects are salient, attracting the observers' attention (Just & Carpenter, 1976). The saccade amplitude is metric that shows the salience or the importance of the visual stimulus. When presented with a salient object, the amplitude tends to be larger (Goldberg & Wichansky, 2003) (as cited in Krassanakis & Filippakopoulou, 2023).

4.5.3 Grayscale statistical heatmaps generation

Another effective way to visualize data, especially gaze data, is grayscale statistical heatmaps. Such heatmaps are a powerful visual tool for understanding visual behavior, because by mapping the distribution of fixations and the duration of each fixation, the areas with the most attention are revealed. In addition, visualizing aggregated gaze data, aids in the identification of common patterns and trends, but also reveals the salient locations, as they incorporate information such as the duration of gaze within specific regions, and thus how long the participants gaze on particular areas. Grayscale

Chapter 4 – Experimental study design and implementation

statistical heatmaps are visually appealing, effectively communicating findings, providing a clear summary of gaze data. Consequently, grayscale statistical heatmaps provide an objective representation of gaze behavior, reducing the reliance on subjective interpretations, as they offer data – driven visualization, quantitatively analyzed.

In this experimental study, grayscale statistical heatmaps are used to present the salient objects of the stimuli observed by the participants. The EyeMMV (Eye Movements Metrics & Visualizations) toolbox (Krassanakis et al., 2014), used for post experimental eye movement analysis, contains a function for heatmap generation. The function implemented is called “heatmap generator”, and a modified version is utilized to export grayscale statistical heatmaps. Five parameters are defined: 1) Spacing between heatmap cells, 2) Maximum horizontal value, 3) Maximum vertical value, 4) Standard deviation (σ) for Gaussian kernel, 5) Size of the Gaussian kernel for the heatmap smoothing. The parameter values are referred in the next paragraph.

Processing the raw gaze data and selecting the parameters fit for this experimental study, the function is implemented using the scripting language of MATLAB, the same language used for the development of the toolbox. The raw gaze data report contains all the participants’ records. The process to be made, is filtering the data and dividing them in separate files, regarding the participant and the data recorded from every visual stimulus observed by the participant. Consequently, $N \times M$ files are created with a specific format (e.g., [p01_Alexandroupoli_16z_Bing.txt]), with N being the total number of the participants and M the total number of the visual stimuli. The exported files contain eight columns: 1) Trial ID, 2) Timestamp, 3) Left gaze X, 4) Left gaze Y, 5) Right gaze X, 6) Right gaze Y, 7) Average gaze X, 8) Average gaze Y. Grayscale statistical heatmaps visualize aggregate data, hence the exported files need to be merged with respect to the visual stimulus. The files to be used are the same number as the visual stimuli used, meaning seventy-five, containing the cumulative data with a specific format (e.g., [Alexandroupoli_16z_Bing.txt]). The records needed for the EyeMMV function are the columns 7 and 8, meaning the average gaze coordinates. As for the rest of the parameters, Gaussian parameters can be calculated by considering factors such as the average distance between the participant and the display screen and the dimensions of the display screen. If the visual angle is around 1 degree (f) and the average distance from the display screen (D) is 600mm, knowing that the distance observed on the display screen (S) can be calculated by the function $S = D * \tan(f)$, converting the result from millimeters to pixels, the σ value is calculated approximately to 38 pixels. The Gaussian kernel equals $6 * \sigma$ value, and the maximum vertical and horizontal values are the resolution of the display screen, 1920 pixels horizontal and 1080 pixels vertical. All pixel values

Chapter 4 – Experimental study design and implementation

are normalized within the range of 0 to 255. Therefore, 8-bit grayscale images are generated with each intensity contained, representing the probability of occurrence of each point. The higher the value (255), the higher the probability of occurrence.

The salient locations are visualized with the grayscale statistical heatmaps, in which every pixel value is normalized, thus comparison can be made to determine quantitative and qualitative differences. The heatmaps produced with the EyeLink® Data Viewer, although visualizing the areas with the most attention for each trial, cannot be compared because the visualized data are not normalized on the same scale, which is the fixation duration. Therefore, the comparisons aim to disclose differences on the attention of the participants whether there is a different cartographic background derived from a different service, or a different zoom level. Using MATLAB, a script is created to visualize the differences. The cartographic backgrounds used as base maps, overlaid with the heatmap comparison image with a transparency of 60%, are the resulted output. The pixel values are within the range of -255 to 255, with values close to 255 being the salient objects of the first heatmap, values near 0 being salient objects on both grayscale statistical heatmaps and values near -255 being salient objects of the second heatmap. Sixty comparisons are made in total, in between the services of each cartographic background, resulting to 4 pairs of comparison per image.

To study the salient objects shown on grayscale statistical heatmaps, the visual stimulus needs to be presented along with the its corresponding statistical grayscale heatmap. Overlaying the grayscale statistical heatmaps on the visual stimuli achieves the visualization of the salient locations with the characteristics of the cartographic background. Observing the areas with the most attention, with the information of the visualized objects on the cartographic background, aids on the qualitative interpretations of the visual stimuli. A script is written using MATLAB, producing seventy-five images with the grayscale statistical heatmaps overlaid on top of the cartographic backgrounds with a transparency of 60%.

Chapter 5 : Results

5.1 Indices analysis

The analysis conducted in Microsoft Excel refers to three metrics that derive from the key records of an eye tracking experiment. The computed indices have a wide range of magnitudes. There are two ways to be compared, either to create separate graphs for each case, or normalize them and compare them all together in a graph. The method of normalization chosen is the log transformation, using the formula $\log(x)$ to base 1000, where x is the value to analyze, hence applying logarithmic function to every value (logarithmic scale). Compressing large values and expanding small values, a single graph can be created containing all values, facilitating the comparisons. The values decided to be analyzed are the median values for each case (median fixation duration, median number of fixations, median saccade amplitude).

Firstly, normalizing the values with the logarithmic formula, sets a new range from 0 to 1. The radar diagram created, depicted in Figure 5.1 visualizes a rather uniform result for all 75 visual stimuli. It seems that the visual stimuli are similar to one another, and for that reason the values of median fixation duration are all equal to approximately 0.8, the values of median number of fixations are equal to 0.3 and the values of median saccade amplitude are equal to 0.2. These results imply that the visual scenes observed require the same levels of visual attention. This aspect needs further examination.

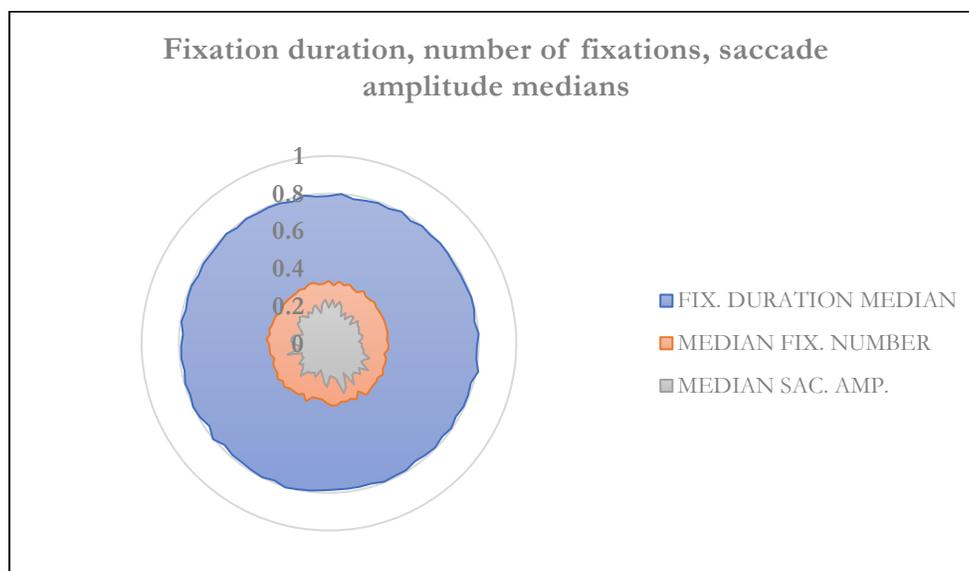


Figure 5.1. Fixation duration, number of fixations and saccade amplitude medians for all visual stimuli, normalized in the range from 0 to 1 with the logarithmic method, visualized in a radar diagram.

To examine if there is, indeed, uniformity on the visual scenes, the analysis is conducted separately for every online map service and zoom level. A figure is created for every online service, for every case, resulting to fifteen figures. Five figures for the fixation duration, five figures for the number of fixations and five figures for saccade amplitude.

5.1.1 Median fixation duration

Visualizing the median fixation duration for every online map service, as mentioned, five figures are created. The values on y – axis represent the milliseconds of the fixation duration, with the range from 0 to 300 milliseconds. Median fixation duration is presented in Figure 5.2 for Bing Maps, in Figure 5.3 for ESRI Maps, in Figure 5.4 for Google Maps, in Figure 5.5 for OSM and in Figure 5.6 for Wikimedia. It is clear that all values are near 250 milliseconds, and more specifically ranging from 212 to 272 milliseconds. Therefore, the fixation duration on every cartographic background, independently of the service used, is of similar value. According to Rayner (1998), the mean fixation duration for visual search is 275 ms and for scene perception 330 ms.

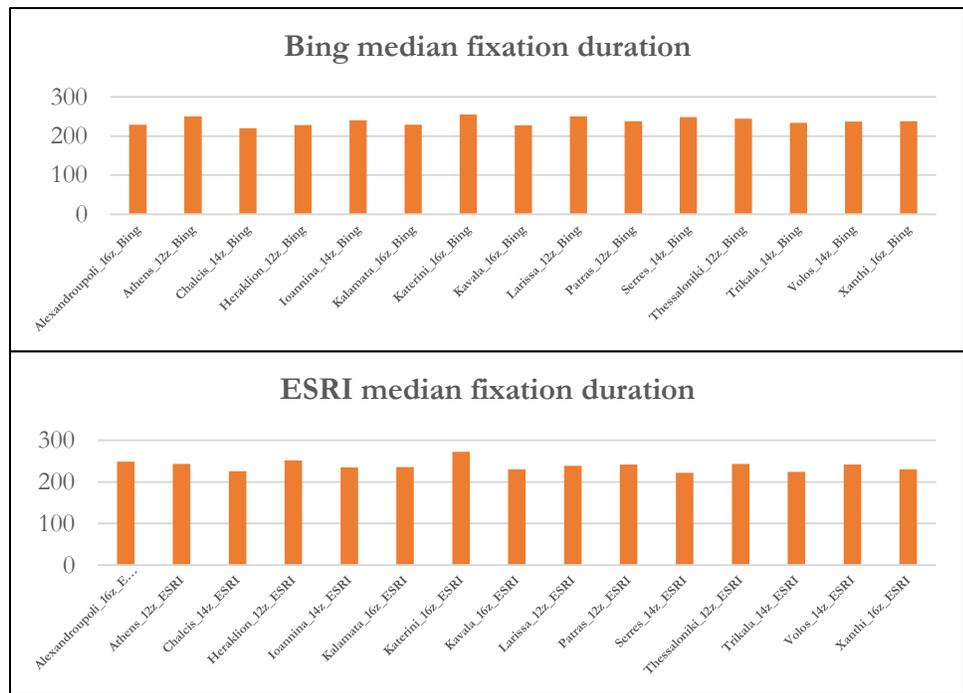


Figure 5.2. Bing median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).

Figure 5.3. ESRI median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).

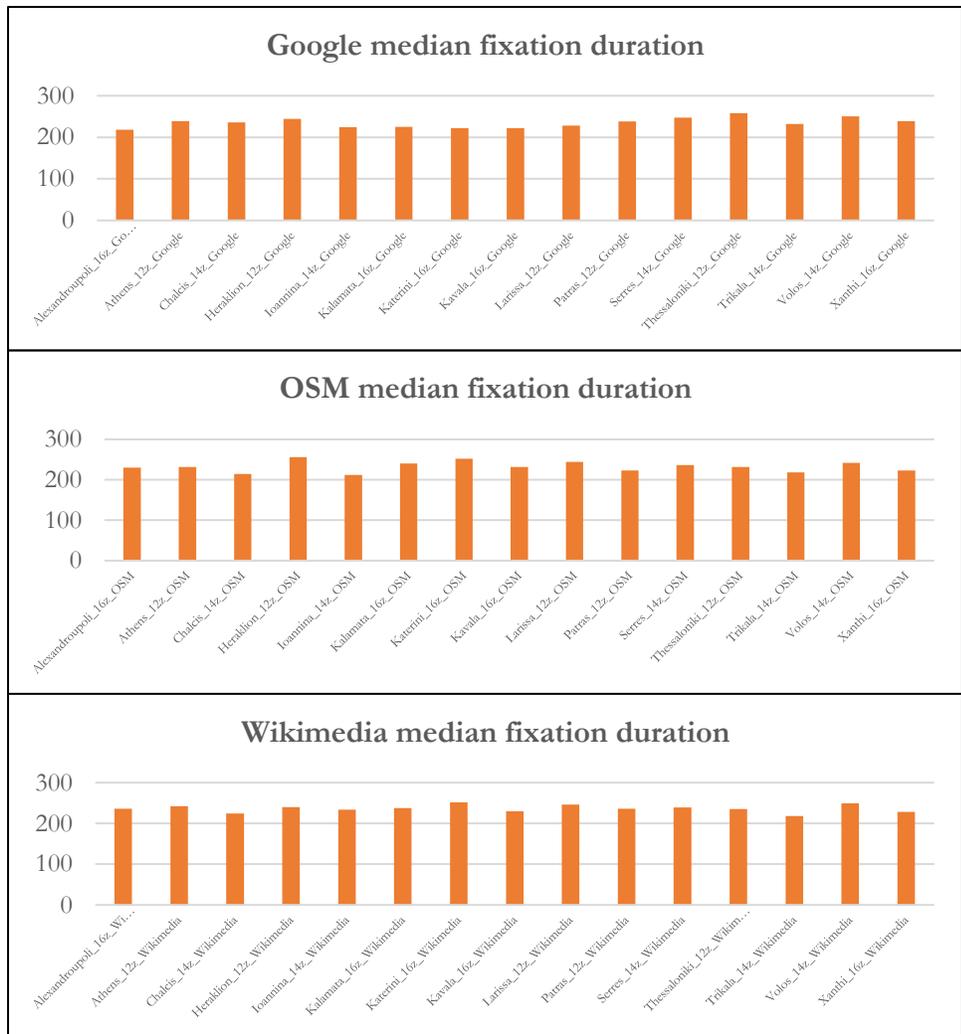


Figure 5.5. Google median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).

Figure 5.4. OSM median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).

Figure 5.6. Wikimedia median fixation duration (x-axis: visual stimuli, y-axis: duration in milliseconds).

5.1.2 Median number of fixations

The figures for the median number of fixations visualize similar results. All values are in range from 8 to 10 fixations per visual stimulus for every online service. The median number of fixations in Figure 5.7 refers to Bing Maps, in Figure 5.8 refers to ESRI Maps, in Figure 5.9 refers to Google Maps, in Figure 5.10 refers to OSM and in Figure 5.11 refers to Wikimedia. Thus, the objects on the cartographic backgrounds the participants explored, are of the same importance and saliency.

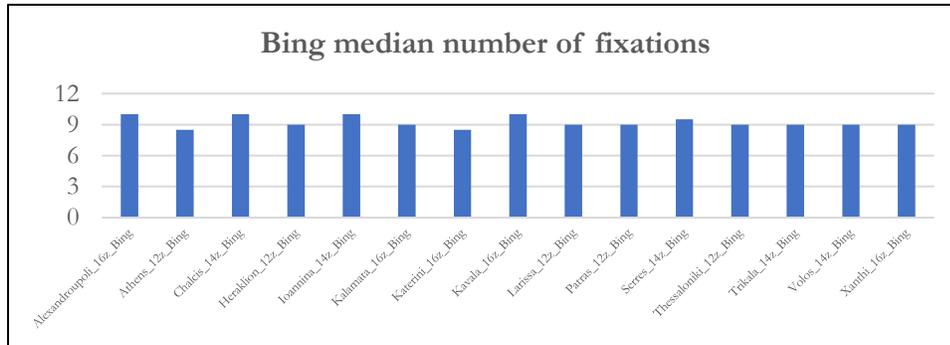


Figure 5.8. Bing median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).

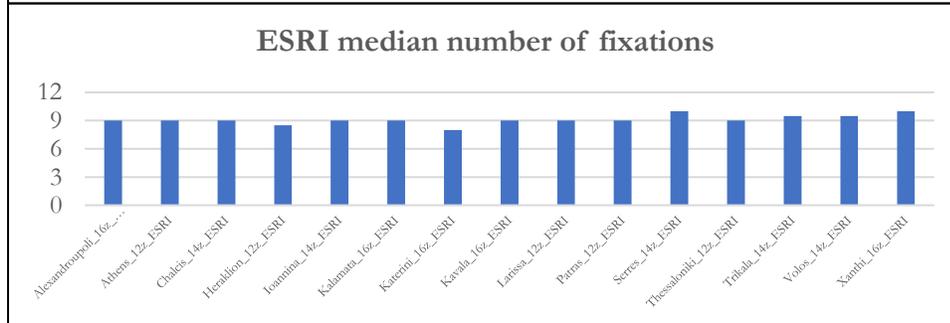


Figure 5.7. ESRI median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).

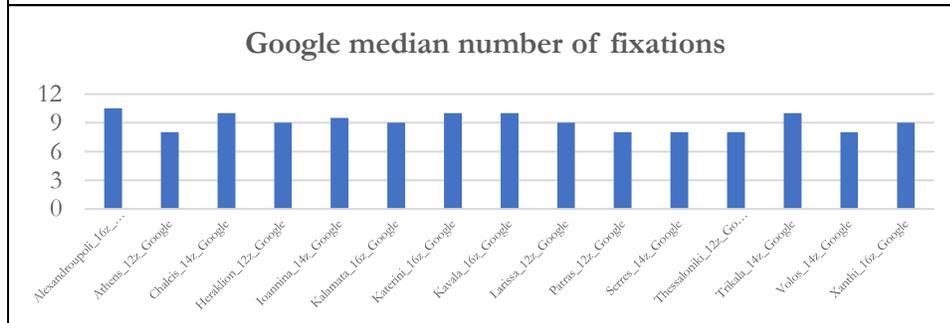


Figure 5.9. Google median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).

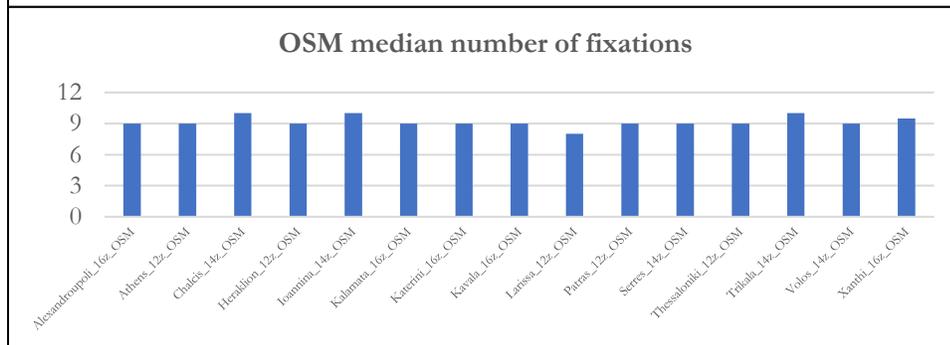


Figure 5.10. OSM median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).

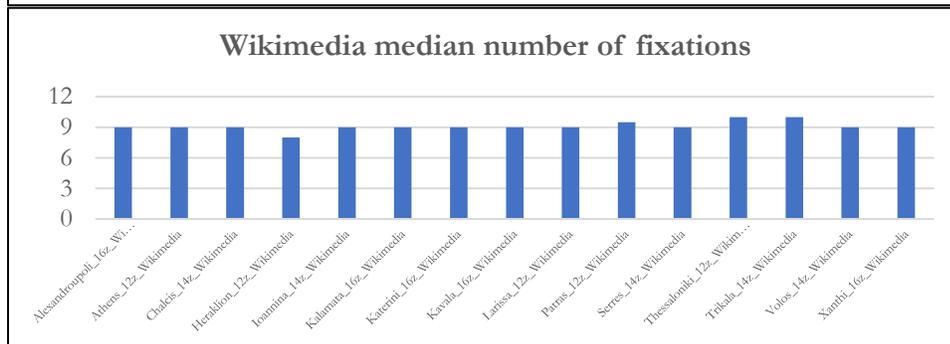


Figure 5.11. Wikimedia median number of fixations (x-axis: visual stimuli, y-axis: number of fixations).

5.1.3 Median saccade amplitude

As for the saccade amplitude the values are similar but the range in medians is greater than the previous two cases. Values range from 2.7 to 6.9, which according to (Rayner, 1998) who approximates the mean saccade amplitude to 3 degrees for visual search and 4 degrees for scene perception, are reasonable for this type of experiment. The range, though, doesn't reveal noticeable differentiations considering there are values up to 39.6 degrees. However, outliers are eliminated utilizing the median value. In Figures 5.12, 5.13, 5.14, 5.15 and 5.16, the median saccade amplitude is depicted for Bing Maps, ESRI Maps, Google Maps, OSM and Wikimedia respectively.

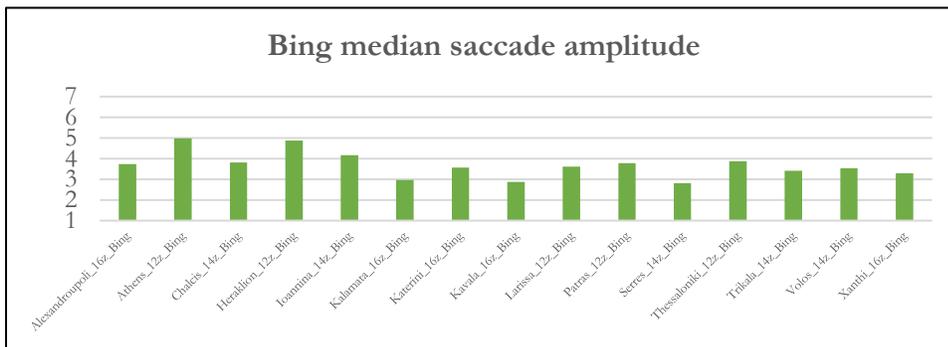


Figure 5.12. Bing median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).

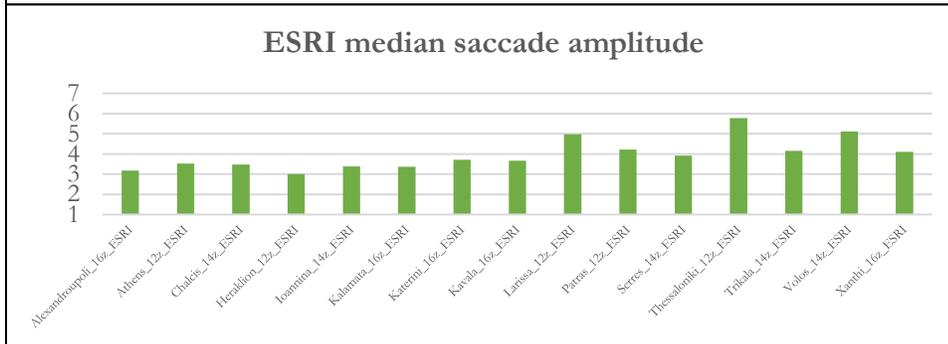


Figure 5.13. ESRI median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).

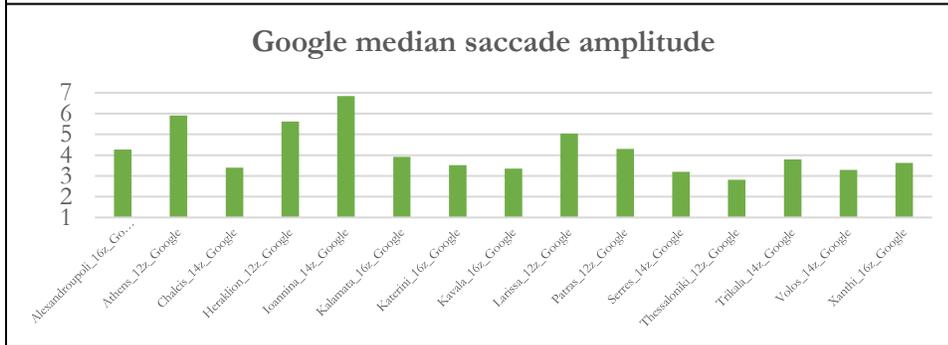


Figure 5.14. Google median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).

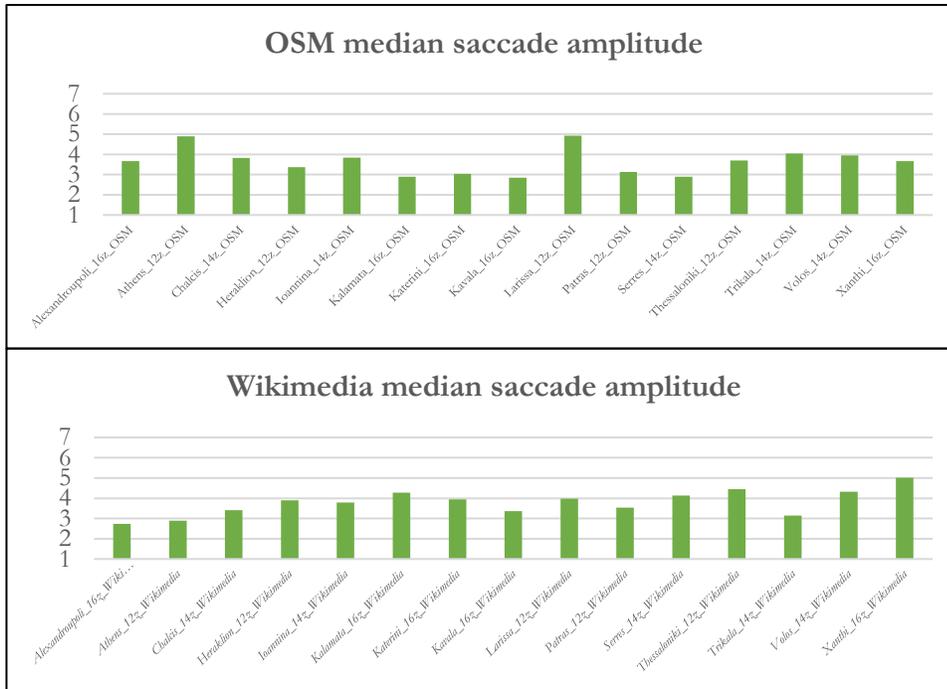


Figure 5.16. OSM median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).

Figure 5.15. Wikimedia median saccade amplitude (x-axis: visual stimuli, y-axis: amplitude in degrees).

5.2 Grayscale statistical heatmaps analysis

The generated grayscale statistical heatmaps visualize the salient objects of the cartographic backgrounds. Cartographic backgrounds from different online map services are visualized following the cartographic design principles, but each service emphasizes on distinct visual variables to make their appearance unique. The objective gaze data of the participants make this aspect clear, because the salient objects in each service provide this type of information. Furthermore, the zoom level plays a role in visual exploration, because the level of abstraction increases as the zoom level decreases. Observing the results, one can understand that in Bing maps, commercial labels are superimposed using red color for their symbolization. ESRI maps represent in detail the elevation relief, emphasize on the location labels with large, bold fonts, and the main roads are in contrast with the neutral colors used in the rest of the map, as they are symbolized with an orange shaded color. Google maps qualitatively differentiate labels using the color visual variable, but are all equal in size. OSM visualize labels in different colors and sizes, while linear characteristics being depicted using both visual variables of color and size. Wikimedia maps emphasize the visualization of the main road net and urban location labels independently of their size, with big cities and smaller towns having the same size labels. Nevertheless, these insights are in general, therefore the examination is conducted for different zoom levels and different cities too. Table 5.1 contains all heatmaps produced from Data

Viewer, grayscale statistical heatmaps that generated implementing EyeMMV toolbox and the grayscale statistical heatmaps overlaid on top of the visual stimuli.

Alexandroupoli, 16z

Bing Maps presents a detailed depiction of the city, showcasing numerous labels. The salient objects encompass labels of restaurants, museums, shops, and the city port, which contribute to a comprehensive view. The label of the city name, in ESRI Maps, is the primary salient object. Following this, there are three labels representing smaller-sized areas that hold significance, with a size larger than other "name" labels. Several salient objects are intoned in Google Maps, including orange hued labels denoting shops, cyan hued labels representing museums, and a red hued label for the city's hospital. In addition, shop labels with a blue hue and the yellow hued main road, along with its associated label, possess a relatively lower degree of prominence. Participants focus on the city's hospital label, in OSM, marking it as a significant point of interest. As for Wikimedia Maps, attention is drawn to labels featuring the names of main roads, thereby capturing their importance within the map context.

Athens, 12z

Participants in Bing Maps visual scene, observe labels of suburb areas situated near the city's name label, highlighting them as the most noticeable objects. In ESRI Maps emphasis is given on the city's name label itself, considering it as the primary salient object. Attention is drawn in Google Maps to both the city's name label and nearby labels with a larger font size compared to the rest, creating a sense of prominence. The city's name label is the salient object on the map for OSM too. Wikimedia Maps features two salient objects: a label representing a suburban area located north of the city's label, followed by the city's label itself.

Chalcis, 14z

Labels of restaurants stand out in Bing Maps, making them salient objects. Bold font labels positioned near the city's name label draws attention in ESRI Maps. These bold labels, which possess a higher visual weight, contribute to the salience of specific points of interest. The city's name label and the label of a well-known landmark are both emphasized in Google Maps. The city's name label is the salient object in OSM too. In Wikimedia Maps salient locations are situated near the city's name label, along with other labels displayed with bold fonts.

Heraklion, 12z

The salient location in Bing Maps, ESRI Maps, Google Maps, OSM, and Wikimedia Maps is the area around the city's name label. This area holds prominence on the map, capturing attention and serving as a focal point for users seeking information about Heraklion. Across these platforms, the salience of the city's name label is emphasized, ensuring its visibility and readability.

Ioannina, 14z

Bing Maps showcases prominent restaurant labels and the name label of the main road as salient objects, as the city's name label is not displayed at this zoom level. In ESRI Maps labels with bold fonts near the only water body depicted in the cartographic background, are identified as the salient objects. One of them is the city's name label. Labels of orange and red hues are salient, with the city's name label attracting the greatest intensity in the case of Google Maps. For OSM a different perspective is presented, where the city's name label does not attract attention at this zoom level. Instead, salient objects include the label for the airport, a label within the depicted water body area, a label adjacent to the water body with bold text, and a label representing a nearby town. Similarly, the observation of the city's name label is happening in Wikimedia Maps. The primary salient locations consist of labels with a red hue, displayed in capital letters but in a smaller font size. Additionally, the presence of ship symbols along the coastline of the water body captures the attention of observers.

Kalamata, 16z

Shop labels with red and purple hues, attract the most attention in Bing Maps. Additionally, labels of two main roads adjacent to the shop labels play a significant role in capturing the observer's focus. In ESRI Maps, three labels are salient objects, featuring bold fonts and a larger size compared to the rest of the labels. Interestingly, the main road with a contrasting hue is not observed as a salient object in this map representation. Labels with orange, blue, and pink hues are primary salient objects in Google Maps, positioned in the center of the cartographic background. Additionally, there are salient objects with grey hues that have drawn attention, too. The road network's main road name labels, primarily located in the upper center of the map, is salient object in OSM. Alongside with it, the label for the city's port, the municipality's stadium, and a red hued cross representing a pharmacy serve as salient objects within the map. In Wikimedia Maps, observers located the labels of the names of the main roads positioned in the center of the map as the most prominent and salient objects.

Katerini, 16z

In Bing Maps attention is drawn to red hued labels, particularly one label on the main road, which stands out as visually contrasting to the rest of the map with its purple hue. These red hued labels serve as salient objects within the visual scene. In ESRI Maps, bold font labels near the center of the map are the salient locations, with one of them representing the name of the city. Google Maps has salient objects pink and blue hue labels dispersed throughout the map. OSM directs attention to several salient objects. The municipal gardens, visualized with a green hue on the left side of the map, along with a blue hued parking label in the center, and the name label of the red hued main road located on the right side of the map, have attracted the most attention within the map context. In Wikimedia Maps, participants observed the most, bold font labels with capital letters primarily positioned in the center of the map, along with one label on the right side. These labels, visually distinct with their bold fonts and capital letters, serve as salient objects. Additionally, attention is drawn to the municipal garden depicted as a green hued area on the left side of the map, although with relatively lower intensity. The label on the right side of the map consists of two words, creating a contrast with the rest of the labels.

Kavala, 16z

Bing Maps concentrates red hued labels in the center of the map. The main road, depicted with a purple hue, and its name label, along with the adjacent red hued labels, emerge as the salient objects within the visual scene, capturing the attention of users. In ESRI Maps, observers place emphasis on bold font labels, including the city's name label. These labels, with a larger font size compared to the rest, become the objects most focused on by participants, creating salience in the cartographic background. Google Maps presents a unique perspective, where the salient object of the visual scene is a grey hued label containing the city's name, but referring to a bus station. Interestingly, the participants' visual search is directed towards the city's name label, which is absent from the map at this zoom level. In OSM, a blue hued label of the city's port, located at the bottom of the visual scene, is the salient location. In Wikimedia Maps, the bold font labels attract the attention of the participants; however, the primary salient object is located in the center of the map.

Larissa, 12z

The salient object is uniform in all online map platforms. The city's name label is located in the center of the visual scenes, inside the circular area created by the main road network. Thus, the salient location in all cartographic backgrounds is this label.

Patra, 12z

Bing Maps presents a visual scene with numerous labels in green hue, primarily representing elevation relief and different locations. The salient objects are positioned in the center of the map excerpt, while the city's name label on the left side exhibits lower intensity relative to the other salient objects. The prominence of these central objects captures the attention of users exploring the map. In ESRI Maps, the observers designate labels adjacent to the main road with a contrasting hue as prominent objects. Additionally, the primary salient object is located near the center of the visual scene, providing a focal point for users navigating the map. In Google Maps, three salient objects are emphasized: the city's name label, the label representing a notable landmark, and the label indicating a mountainous area. These objects are considered significant points of interest within the map context. In OSM, discretion is revealed in the intensity values of labels near the main road, highlighting their varying importance. However, the salient locations include a large green area near the center of the visual scene, representing a significant natural feature, and a label situated right next to the main road, characterized by a contrasting red hue. For Wikimedia Maps, the city's name label and location labels near the center of the cartographic background attract the most attention.

Serres, 14z

In Bing Maps, the salient objects are situated in the center of the visual scene, where the city's name label is located. Additionally, red and blue hued labels representing hotels and restaurants contribute to the saliency of this area. The saliency of the city's name label is emphasized in ESRI Maps, as its intensity clearly distinguishes it as the primary object of interest within the visual scene. The salient objects, are highlighted in Google Maps, in the center of the cartographic background. Adjacent to the city's name label, there is a green hued label representing a park and an orange hued label representing a shop. These objects are the prominent features of the map excerpt. Although there may be other objects with increased intensity, their saliency is incomparable to those located in the center. In OSM the city's name label is designated as the salient object, similarly to Wikimedia Maps.

Thessaloniki, 12z

The grayscale statistical heatmaps of this area show a uniformity on the salient objects. In every online map service, the participants have observed the city's name label and a label near the water body in the center of the cartographic background. On top of that, the intensity of the latter object is approximately the same with the city's name label.

Trikala, 14z

In Bing Maps, ESRI Maps and Wikimedia Maps, the salient object is the city's name label. In Google Maps, there are, also, pink hued labels that drew attention and in OSM, an intersection of the main road network has an increased intensity, but not as much as the label of the city name.

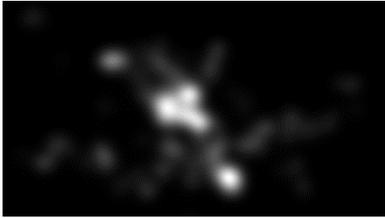
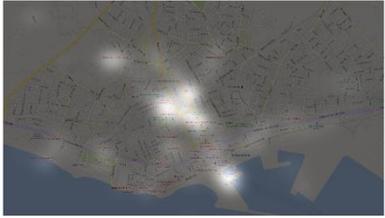
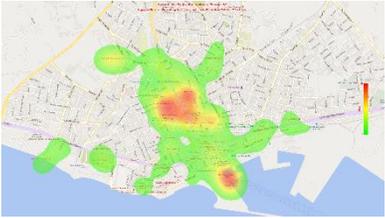
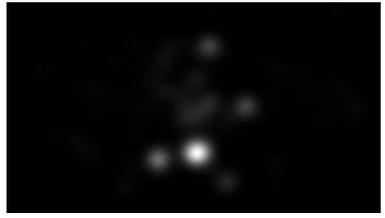
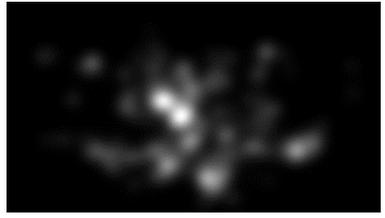
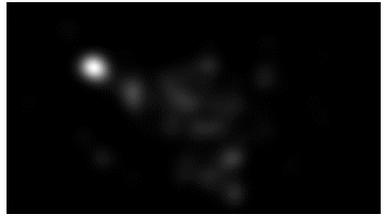
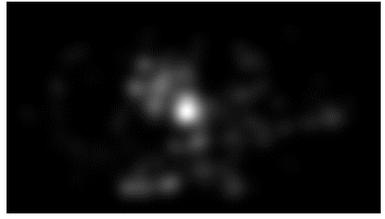
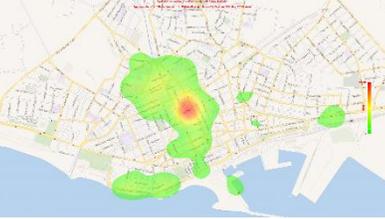
Volos, 14z

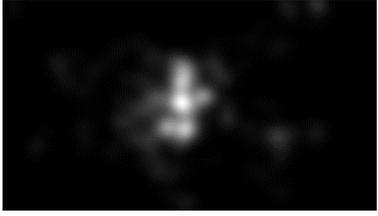
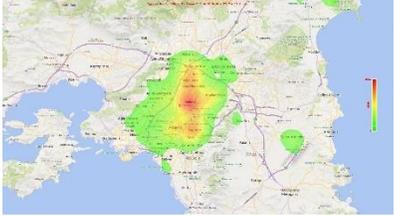
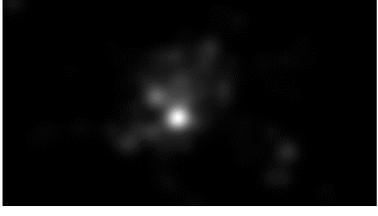
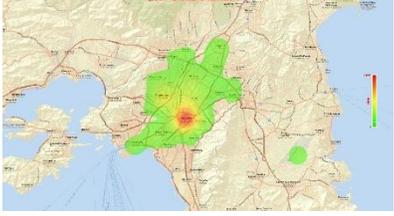
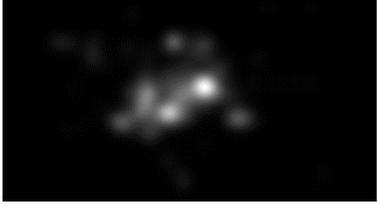
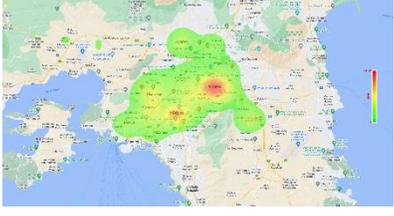
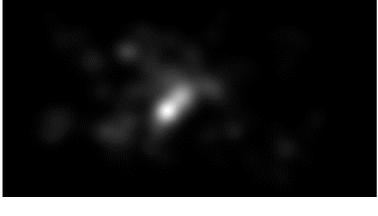
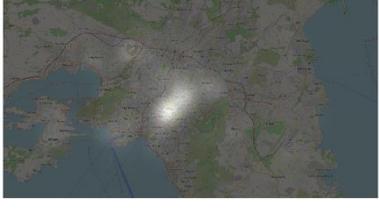
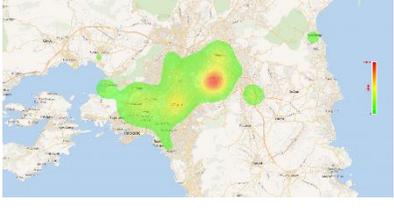
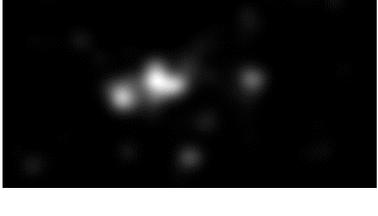
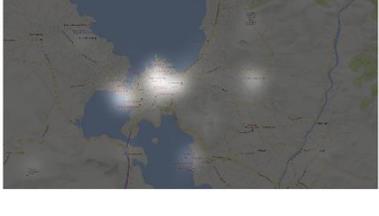
In Bing Maps, the numerous red hued labels, adjacent to the city's name label, serve as salient objects. In ESRI Maps, the absence of the city's name label, made other labels with bold fonts, the salient locations of the map excerpt. In Google Maps, it is clear that the city's name label is the salient object, along with the municipality's name label near it. A location near the city's name label is more prominent, with higher intensity in OSM, even though the font size is the same. However, in Wikimedia Maps, that location is the salient object.

Xanthi, 16z

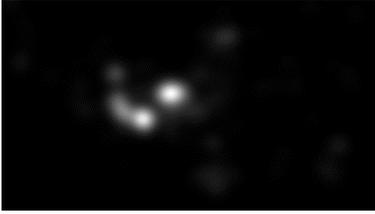
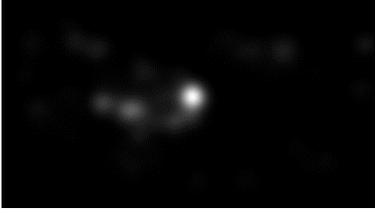
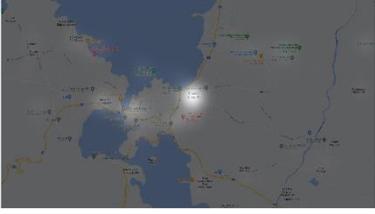
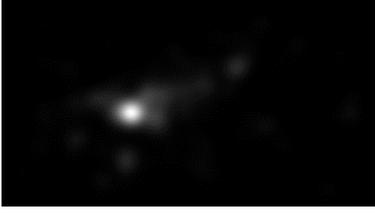
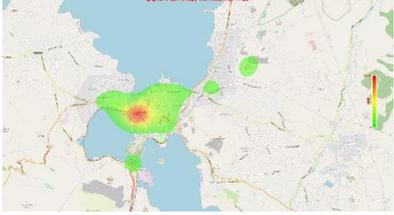
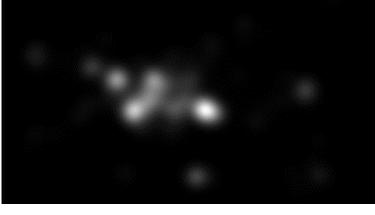
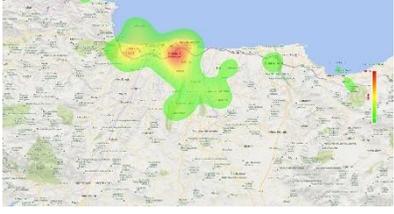
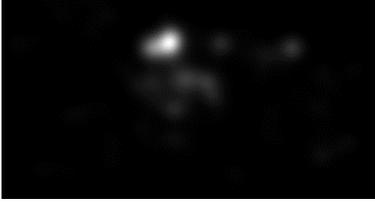
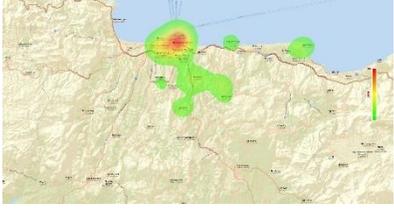
In Bing Maps, red hued labels in the center of the visual scene are prominent, along with the label of the main road on top of the visual scene, which is contrasting with its purple hue the green background of the area located. ESRI Maps have bold font in the labels of the city's name and the suburbs, resulting being the salient objects. In Google Maps, the city's name label is absent, but blue hued labels located where it is in other service, attracted attention. A red hued label on the left of the visual scene in OSM, attracted attention with its contrasting visualization. Also, the center of the map excerpt is a salient location, where the city's name label is usually located, but now is absent. Wikimedia Maps, in this zoom level, visualize the label of suburban area with a font larger than the rest of the labels, making it salient object according to the grayscale statistical heatmaps.

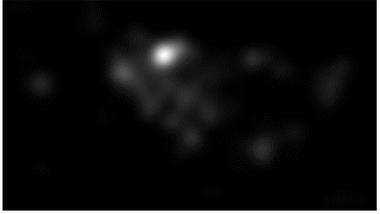
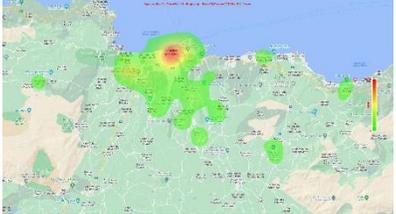
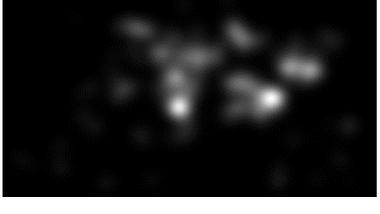
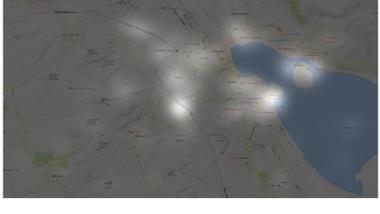
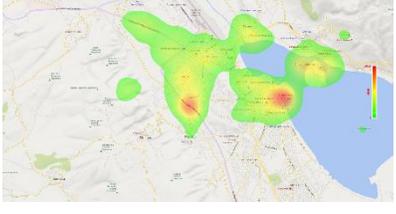
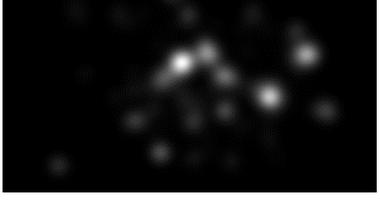
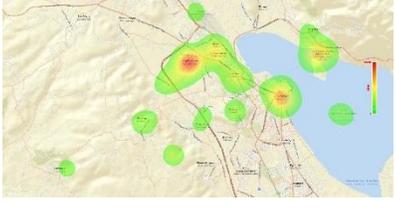
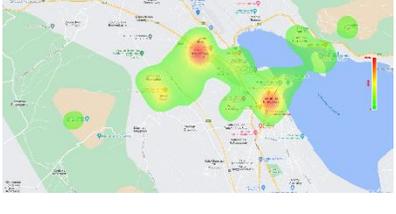
Table 5.1. Generated grayscale statistical heatmaps with heatmap generator script implementing EyeMMV toolbox (left), overlaid cartographic backgrounds with generated grayscale statistical heatmaps (middle), heatmaps generated from Eyelink Data Viewer (right).

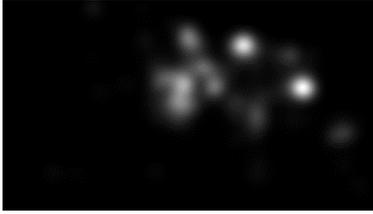
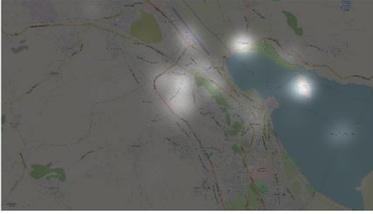
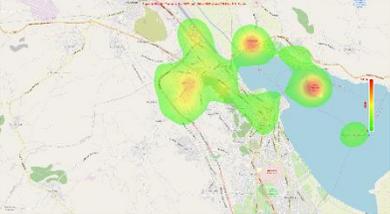
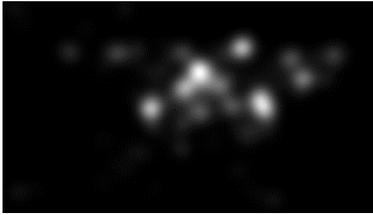
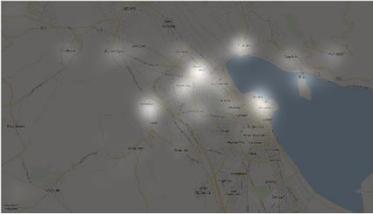
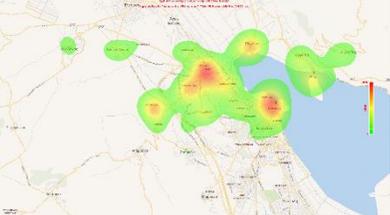
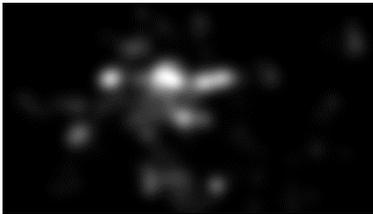
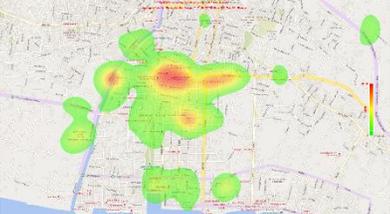
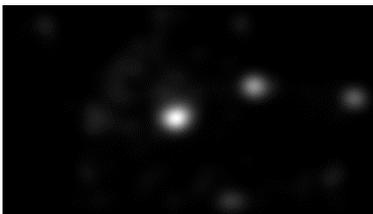
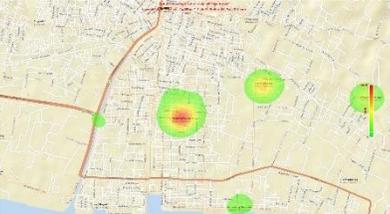
Name	Heatmap Generator	Overlaid Heatmaps	Data Viewer Heatmaps
Alexandroupoli 16z Bing			
Alexandroupoli 16z ESRI			
Alexandroupoli 16z Google			
Alexandroupoli 16z OSM			
Alexandroupoli 16z Wikimedia			

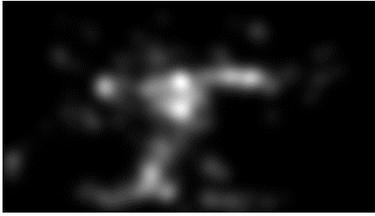
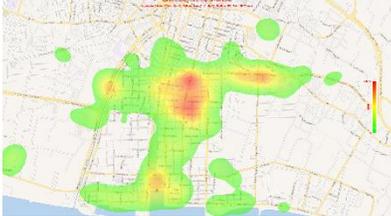
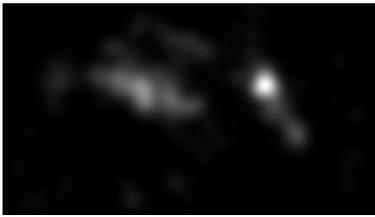
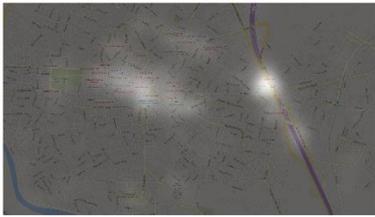
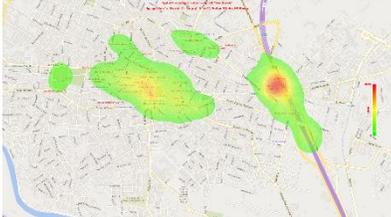
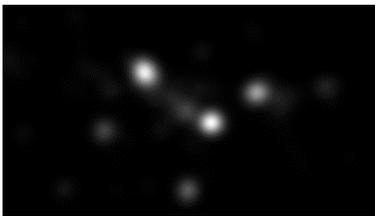
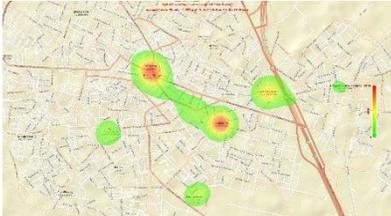
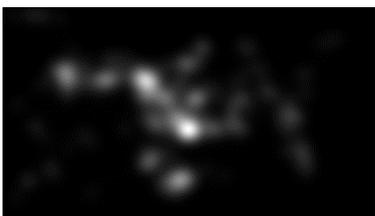
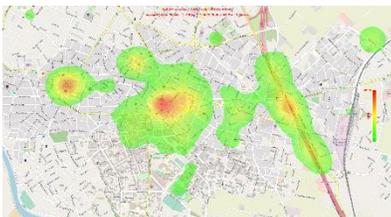
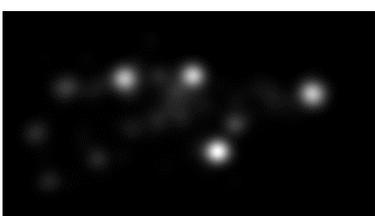
<p>Athens 12z Bing</p>			
<p>Athens 12z ESRI</p>			
<p>Athens 12z Google</p>			
<p>Athens 12z OSM</p>			
<p>Athens 12z Wikimedia</p>			
<p>Chalcis 14z Bing</p>			

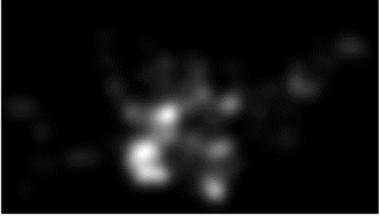
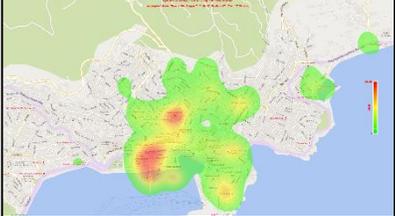
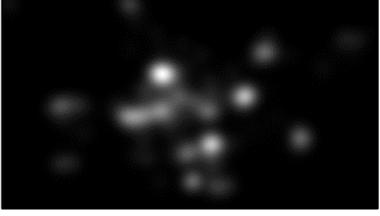
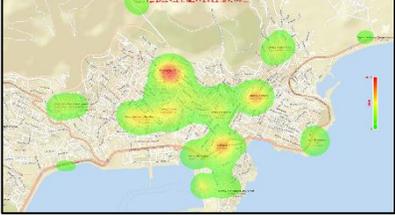
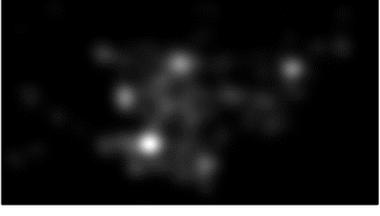
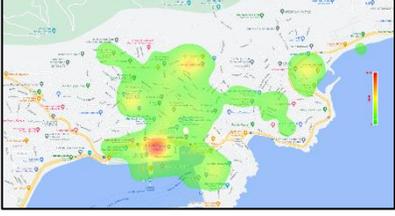
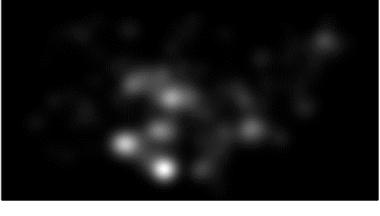
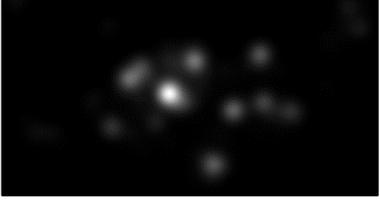
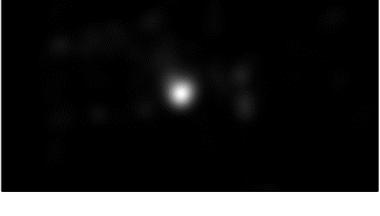
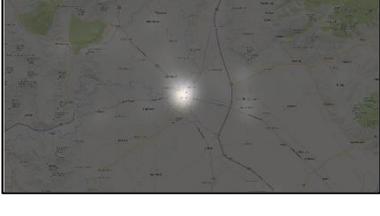
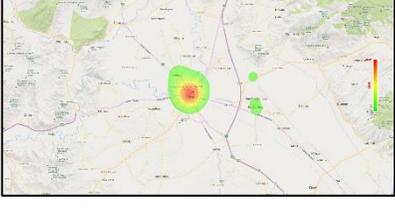
Chapter 5 - Results

<p>Chalcis 14z ESRI</p>			
<p>Chalcis 14z Google</p>			
<p>Chalcis 14z OSM</p>			
<p>Chalcis 14z Wikimedia</p>			
<p>Heraklion 12z Bing</p>			
<p>Heraklion 12z ESRI</p>			

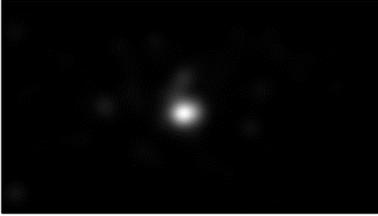
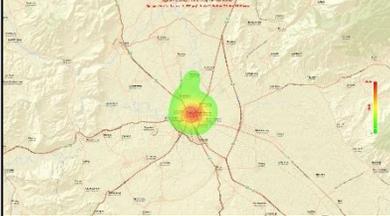
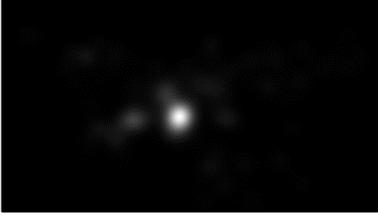
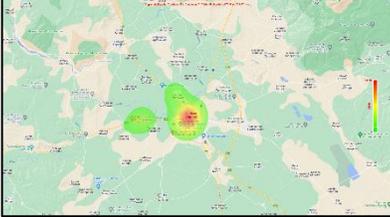
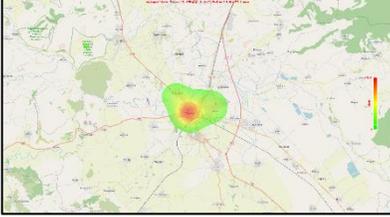
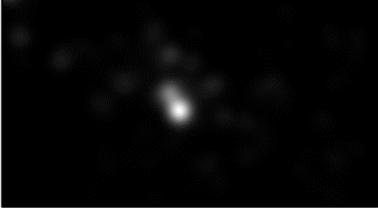
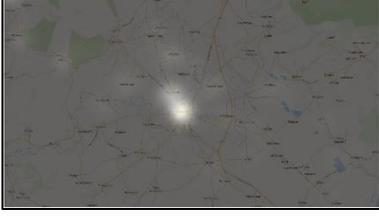
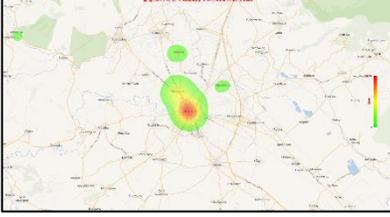
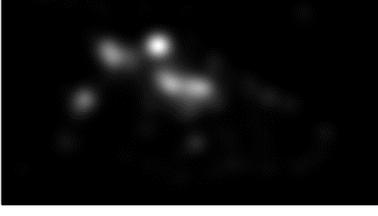
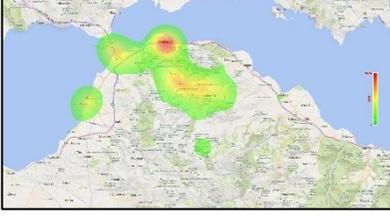
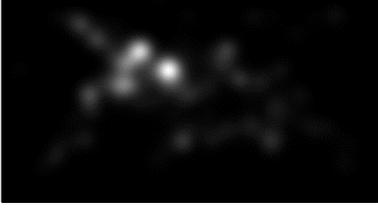
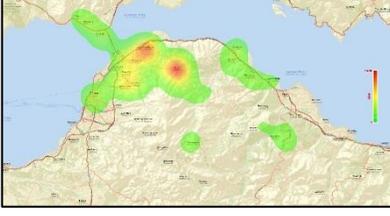
<p>Heraklion 12z Google</p>			
<p>Heraklion 12z OSM</p>			
<p>Heraklion 12z Wikimedia</p>			
<p>Ioannina 14z Bing</p>			
<p>Ioannina 14z ESRI</p>			
<p>Ioannina 14z Google</p>			

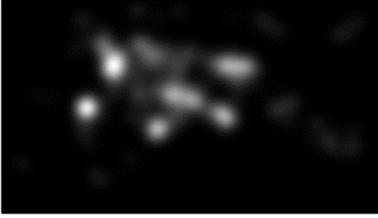
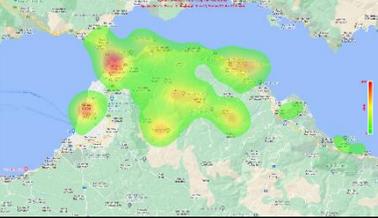
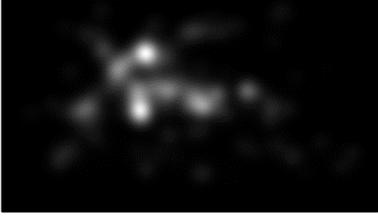
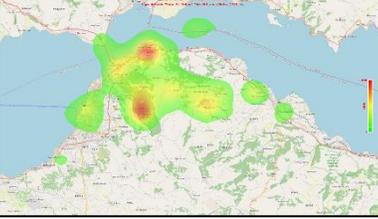
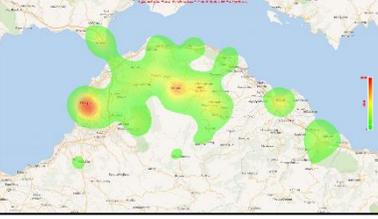
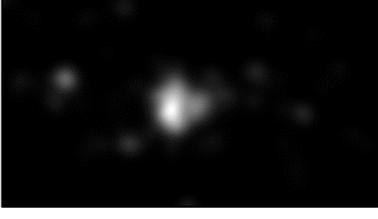
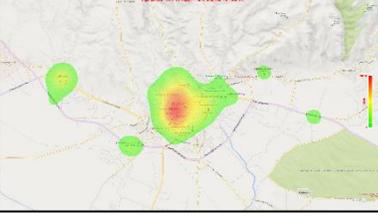
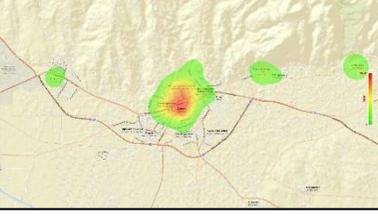
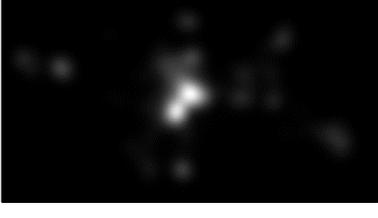
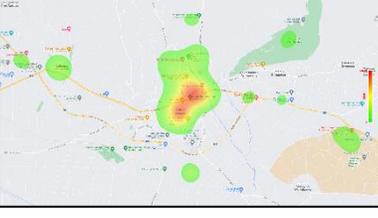
<p>Ioannina 14z OSM</p>			
<p>Ioannina 14z Wikimedia</p>			
<p>Kalamata 16z Bing</p>			
<p>Kalamata 16z ESRI</p>			
<p>Kalamata 16z Google</p>			
<p>Kalamata 16z OSM</p>			

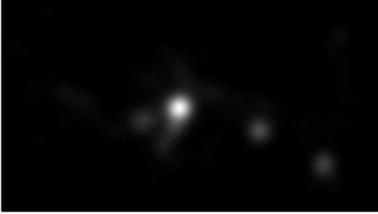
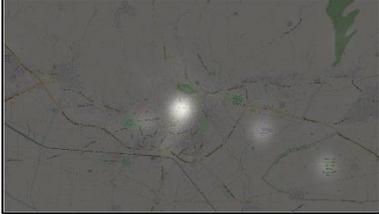
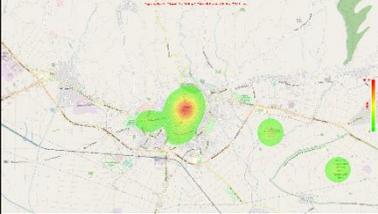
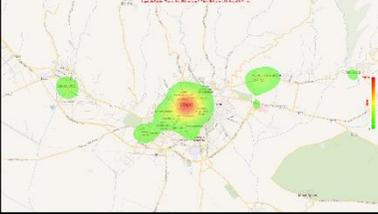
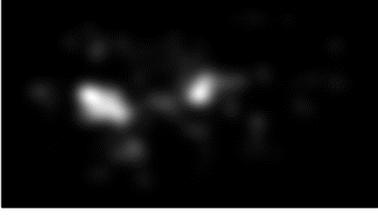
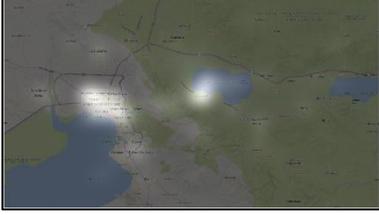
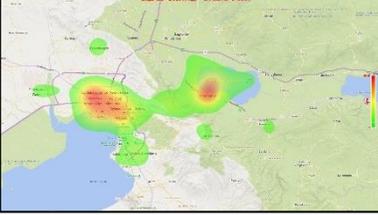
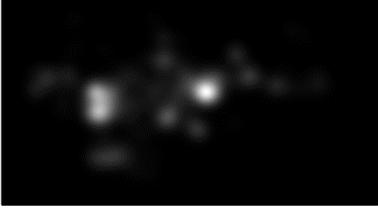
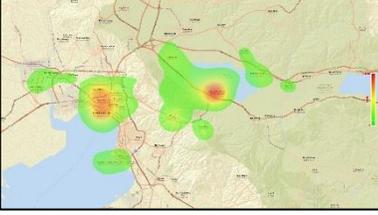
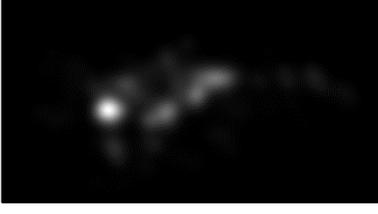
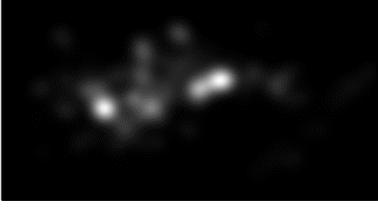
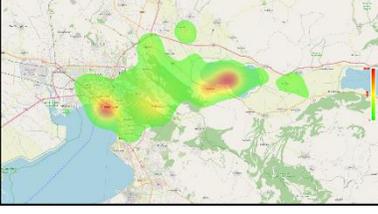
<p>Kalamata 16z Wikimedia</p>			
<p>Katerini 16z Bing</p>			
<p>Katerini 16z ESRI</p>			
<p>Katerini 16z Google</p>			
<p>Katerini 16z OSM</p>			
<p>Katerini 16z Wikimedia</p>			

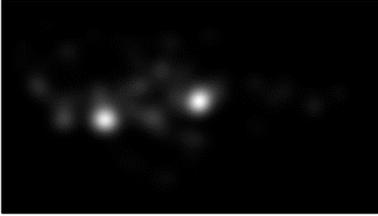
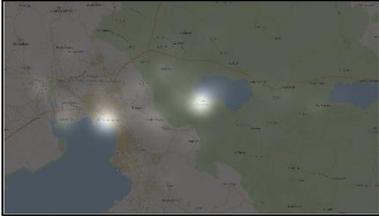
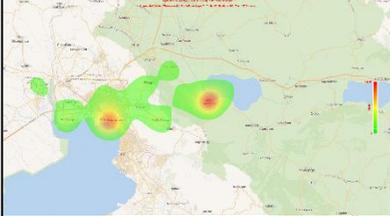
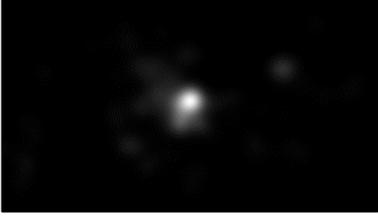
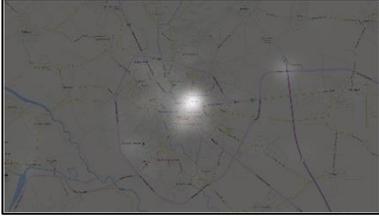
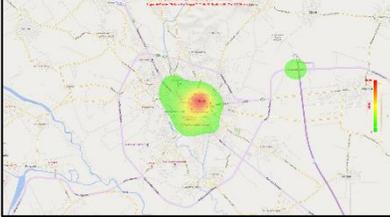
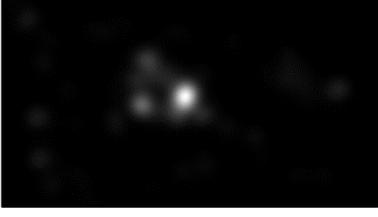
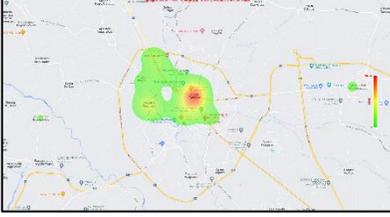
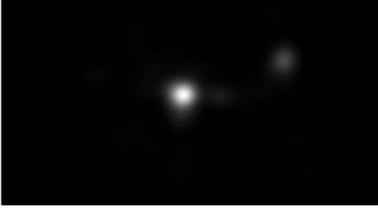
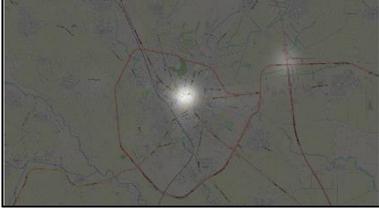
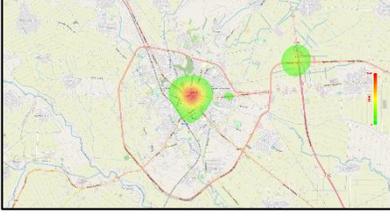
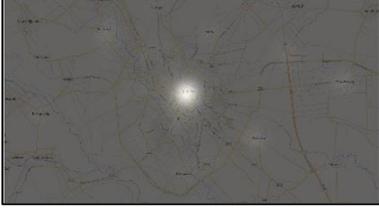
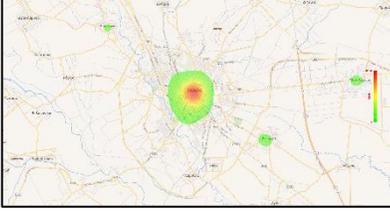
<p>Kavala 16z Bing</p>			
<p>Kavala 16z ESRI</p>			
<p>Kavala 16z Google</p>			
<p>Kavala 16z OSM</p>			
<p>Kavala 16z Wikimedia</p>			
<p>Larissa 12z Bing</p>			

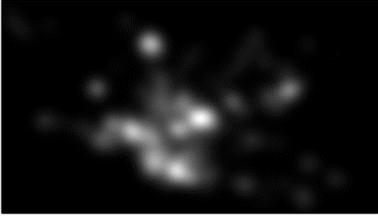
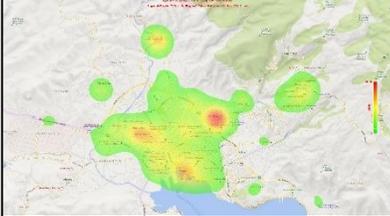
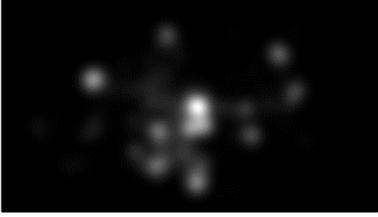
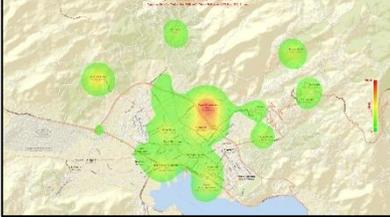
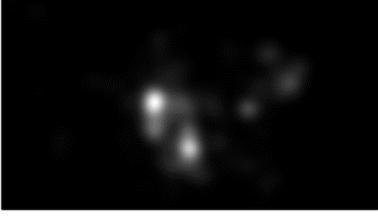
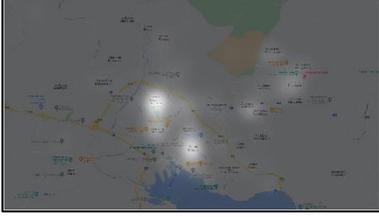
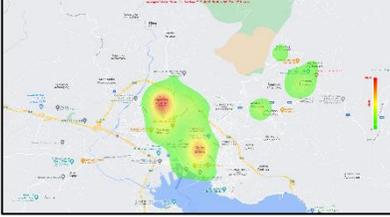
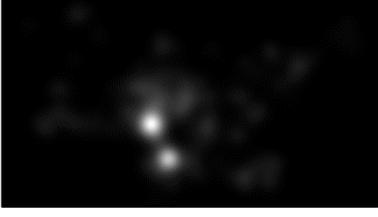
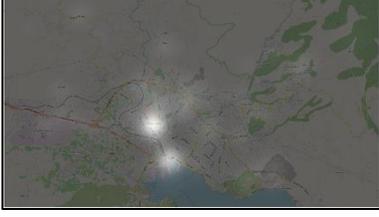
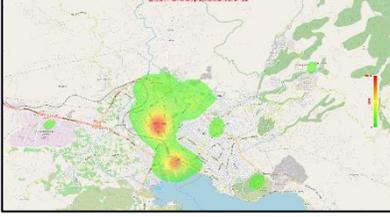
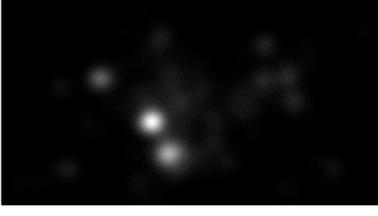
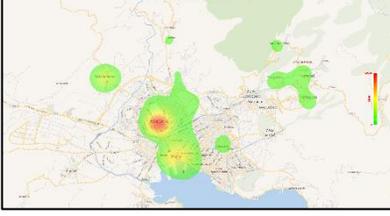
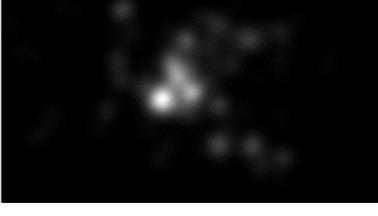
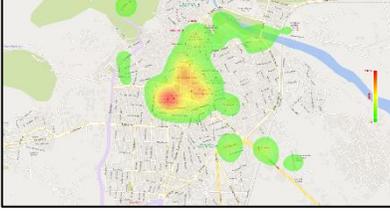
Chapter 5 - Results

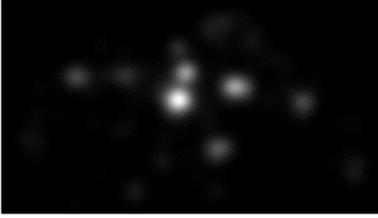
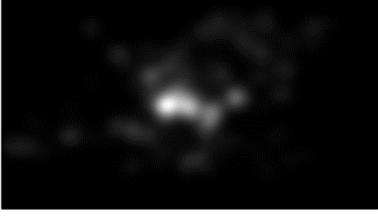
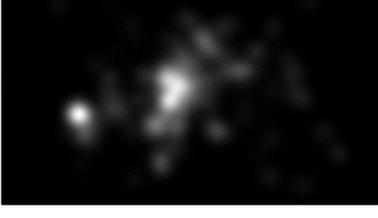
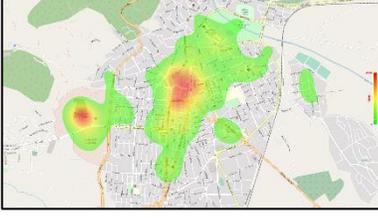
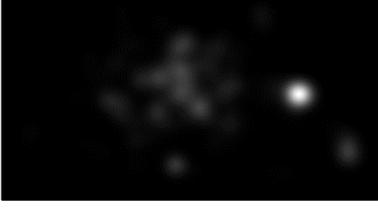
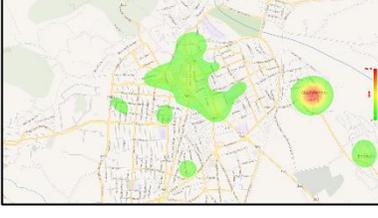
<p>Larissa 12z ESRI</p>			
<p>Larissa 12z Google</p>			
<p>Larissa 12z OSM</p>			
<p>Larissa 12z Wikimedia</p>			
<p>Patras 12z Bing</p>			
<p>Patras 12z ESRI</p>			

<p>Patras 12z Google</p>			
<p>Patras 12z OSM</p>			
<p>Patras 12z Wikimedia</p>			
<p>Serres 14z Bing</p>			
<p>Serres 14z ESRI</p>			
<p>Serres 14z Google</p>			

<p>Serres 14z OSM</p>			
<p>Serres 14z Wikimedia</p>			
<p>Thessaloniki 12z Bing</p>			
<p>Thessaloniki 12z ESRI</p>			
<p>Thessaloniki 12z Google</p>			
<p>Thessaloniki 12z OSM</p>			

<p>Thessaloniki 12z Wikimedia</p>			
<p>Trikala 14z Bing</p>			
<p>Trikala 14z ESRI</p>			
<p>Trikala 14z Google</p>			
<p>Trikala 14z OSM</p>			
<p>Trikala 14z Wikimedia</p>			

<p>Volos 14z Bing</p>			
<p>Volos 14z ESRI</p>			
<p>Volos 14z Google</p>			
<p>Volos 14z OSM</p>			
<p>Volos 14z Wikimedia</p>			
<p>Xanthi 16z Bing</p>			

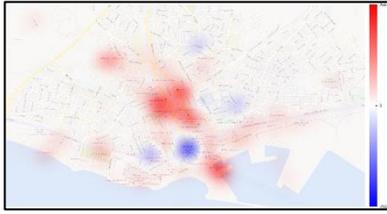
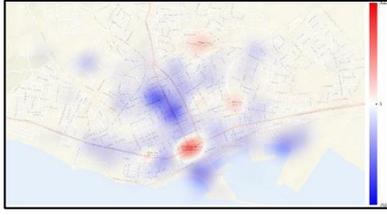
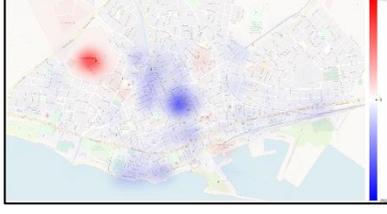
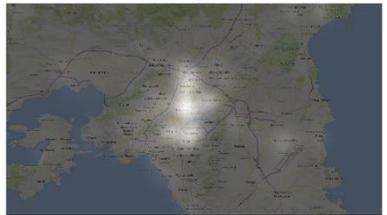
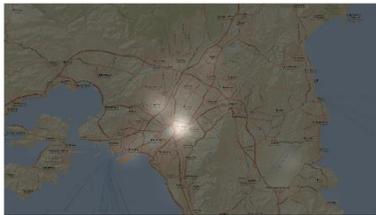
<p>Xanthi 16z ESRI</p>			
<p>Xanthi 16z Google</p>			
<p>Xanthi 16z OSM</p>			
<p>Xanthi 16z Wikimedia</p>			

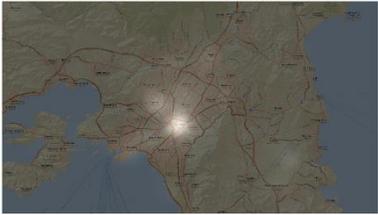
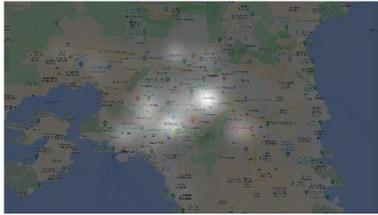
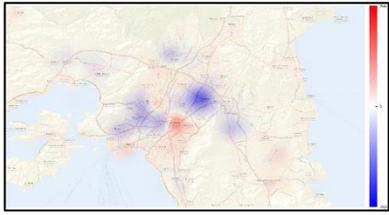
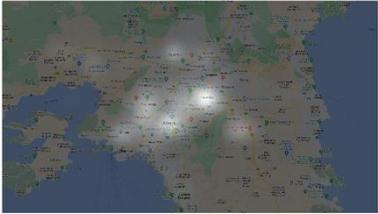
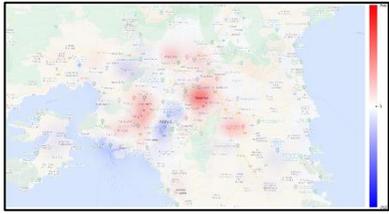
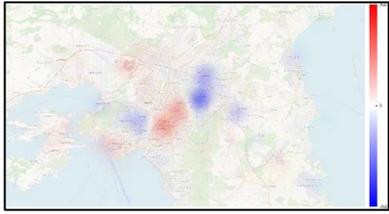
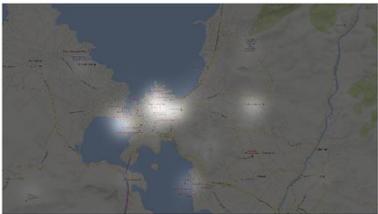
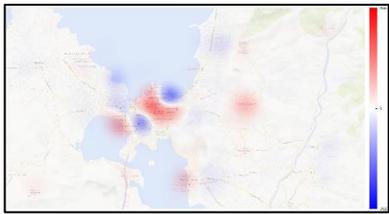
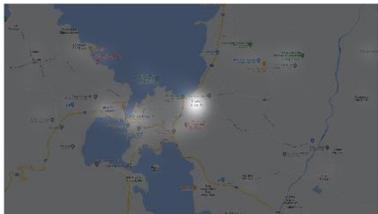
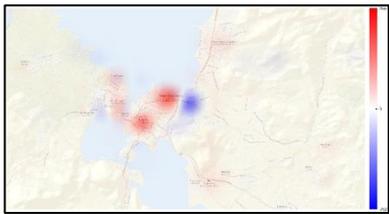
5.3 Grayscale statistical heatmaps comparison

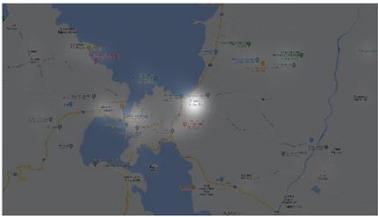
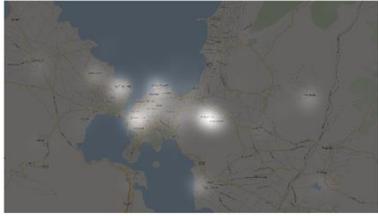
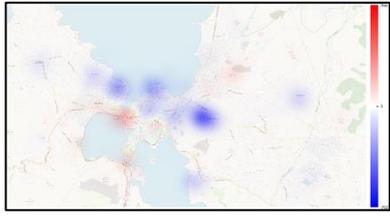
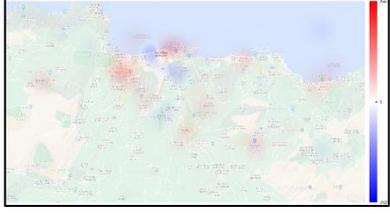
Comparing the grayscale statistical heatmaps is process that provides further understanding and verification concerning the saliency of the objects observed in the grayscale statistical heatmaps analysis, and insight relative to similarities in the cartographic design of the visual scenes. The patterns presented in some of the grayscale statistical heatmaps seem similar to each other, and one can interpretate these results as the existence of uniformity in the visualization of the visual stimuli. Yet, the comparisons prove the opposite. Labels are placed in different locations even if it doesn't seem so, and even if the cartographic backgrounds visualize the same areas, the design principles to achieve visual balance and proper hierarchy in the visual variables used, sometimes are not applied effectively. Consequently, if there would be uniformity in the results, the comparisons would not have that many differences, as the results reveal, and the salient objects would be the same in all visual stimuli.

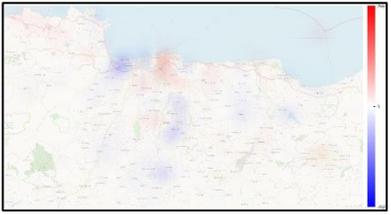
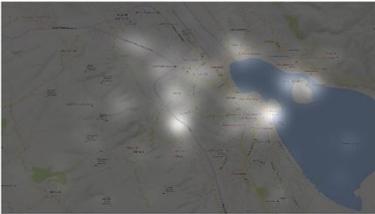
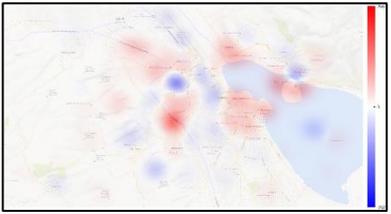
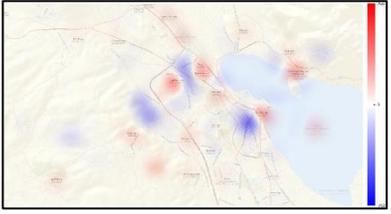
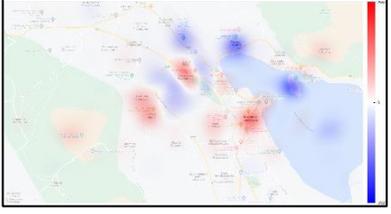
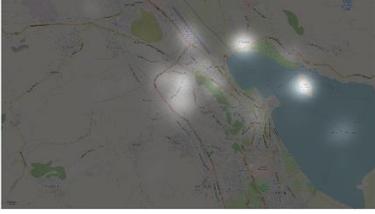
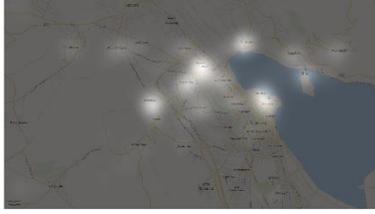
The comparisons were made for each city's cartographic backgrounds, in between the online map services. Bing maps heatmap is compared to ESRI maps heatmap. Then ESRI maps heatmap is compared to Google maps heatmap. Google maps heatmap is compared to OSM heatmap and lastly OSM heatmap is compared to Wikimedia heatmap. This pattern is followed for all cities. There isn't any targeted comparison, because in this stage, emphasis is given to the qualitative differentiations of the visual stimuli. The result is produced by subtracting one heatmap from another. The values are already normalized from 0 value to 255. Visualizing the resulted comparison heatmap, the values of the first heatmap are ranged from 0 to 255 symbolized with the intensity visual variable of red hue, and the values of the second heatmap, ranged from 0 to -255, are symbolized using the intensity visual variable of blue hue. The common values are equal to 0, symbolized with white hue. Table 5.2 contains the comparisons made. The first column is the name of the city, the second column is the subtraction of the grayscale statistical heatmaps, the next two columns present the grayscale statistical heatmaps subtracted from one another and the last column is the result of the subtraction.

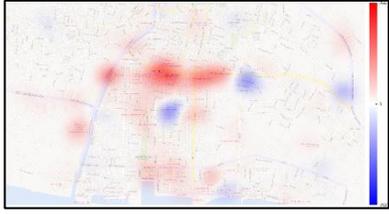
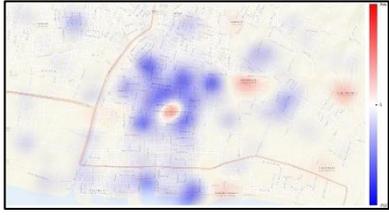
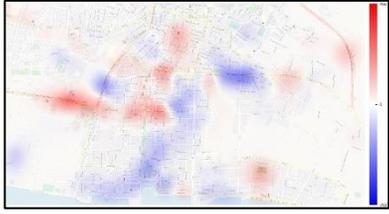
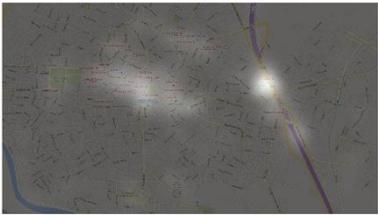
Table 5.2. Grayscale statistical heatmap comparisons for all visual stimuli. Every city has four comparison heatmaps with regard to the online map services that the visual stimuli derive from.

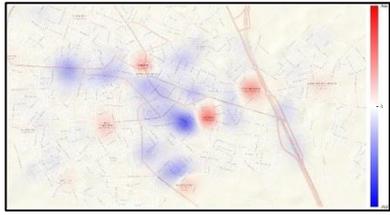
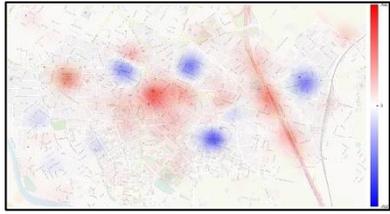
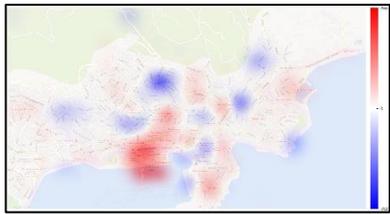
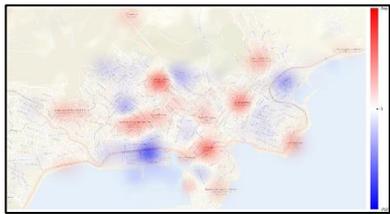
Name	H1-H2	Heatmap 1	Heatmap 2	Comparison Heatmap
Alexandroupoli 16z	Bing - ESRI			
	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			
Athens 12z	Bing - ESRI			

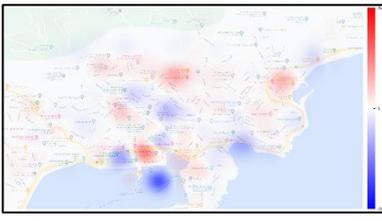
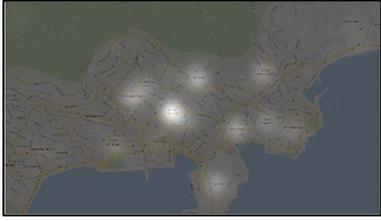
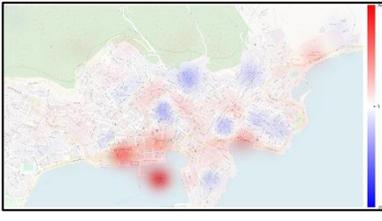
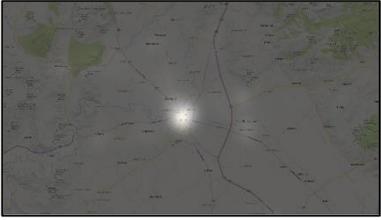
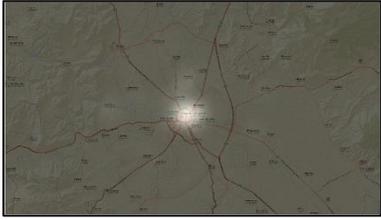
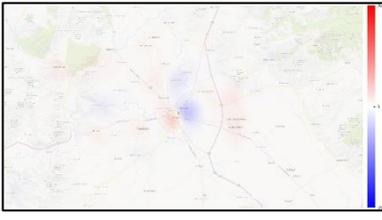
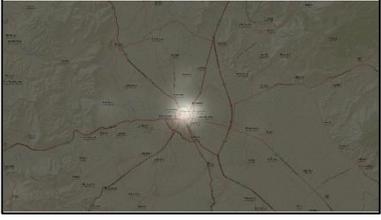
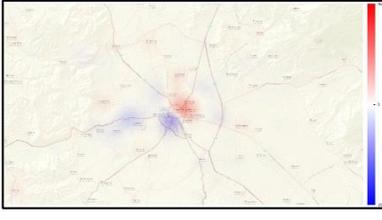
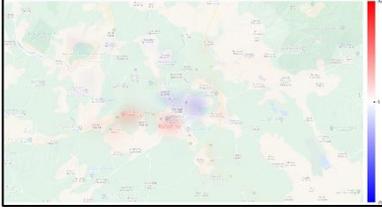
	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			
Chalcis 14z	Bing - ESRI			
	ESRI - Google			

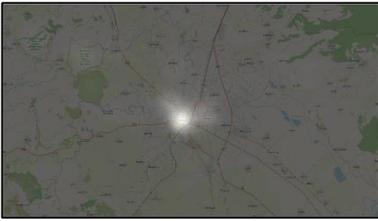
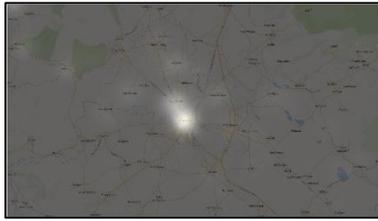
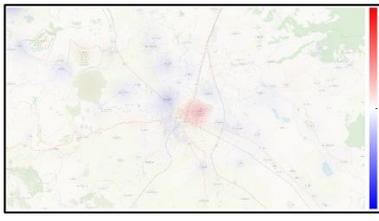
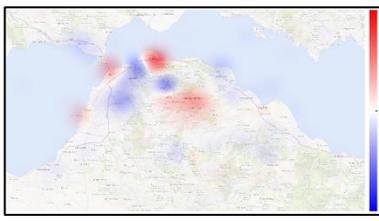
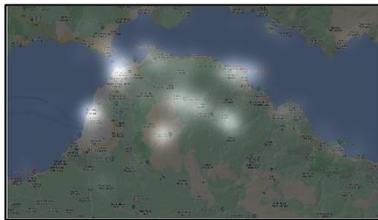
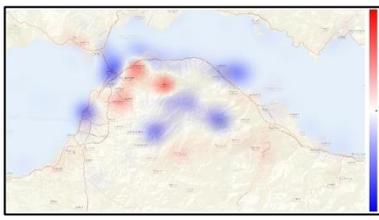
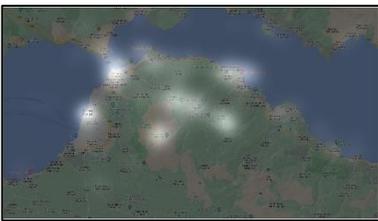
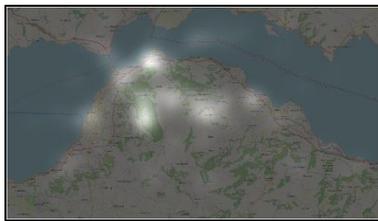
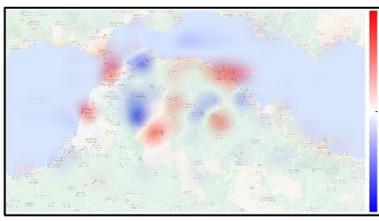
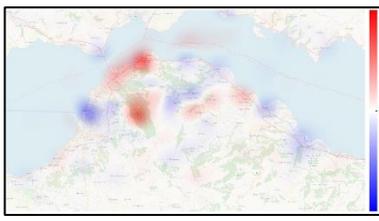
	Google - OSM			
	OSM - Wikimedia			
Heraklion 12z	Bing - ESRI			
	ESRI - Google			
	Google - OSM			

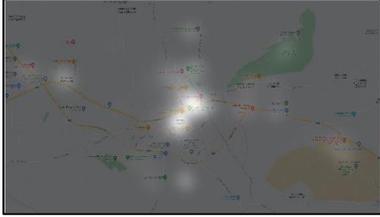
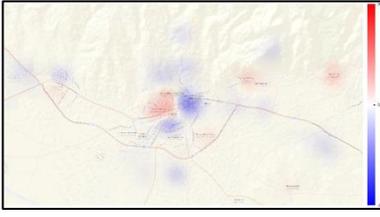
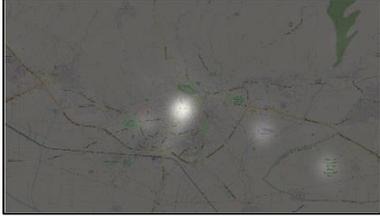
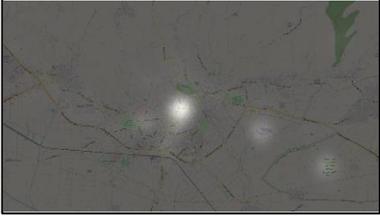
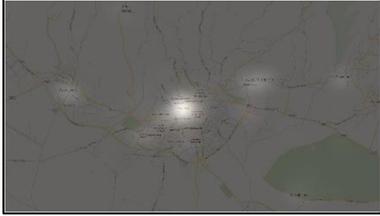
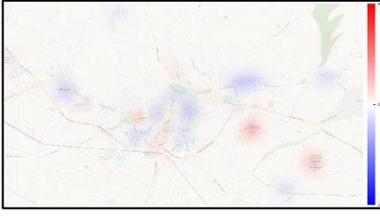
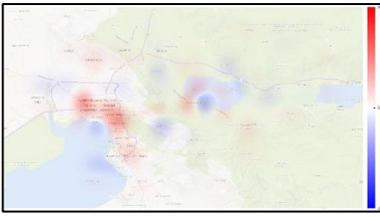
	OSM - Wikimedia			
Ioannina 14z	Bing - ESRI			
	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			

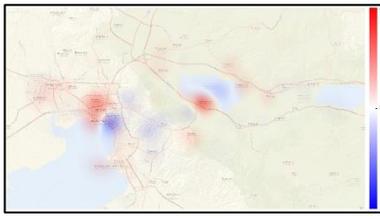
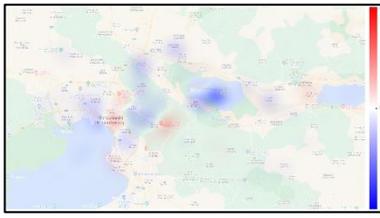
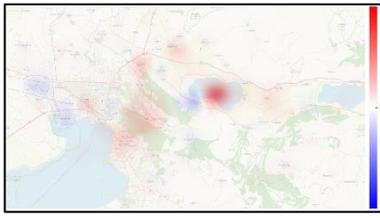
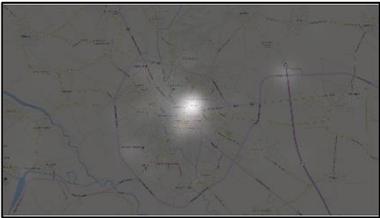
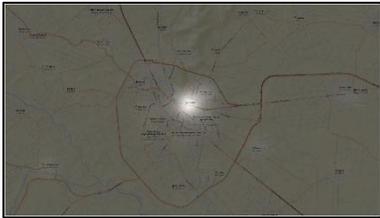
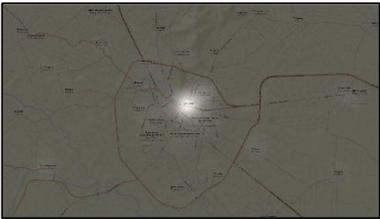
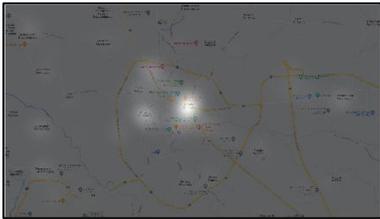
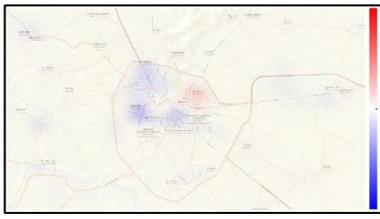
Kalamata 16z	Bing - ESRI			
	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			
Katerini 16z	Bing - ESRI			

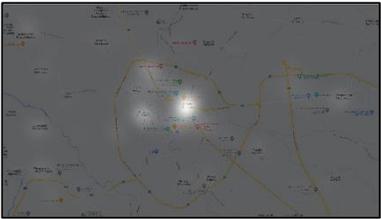
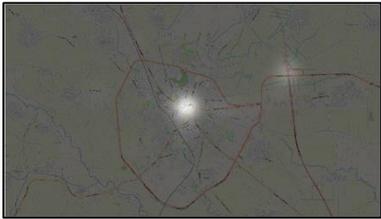
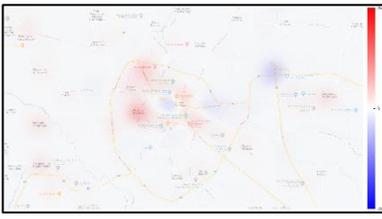
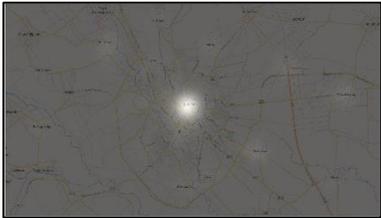
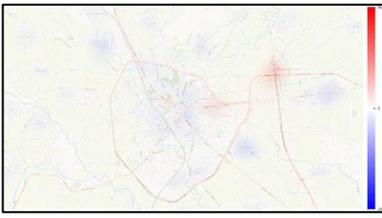
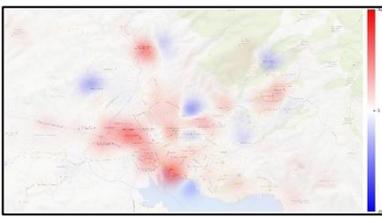
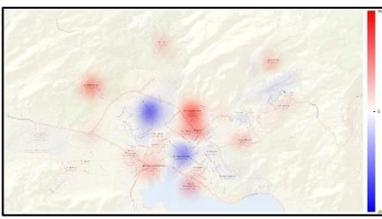
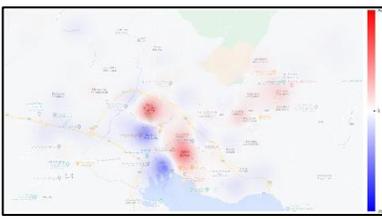
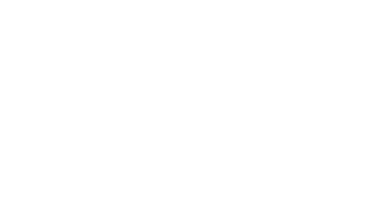
	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			
Kavala 16z	Bing - ESRI			
	ESRI - Google			

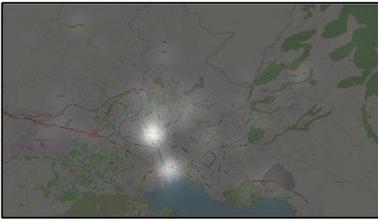
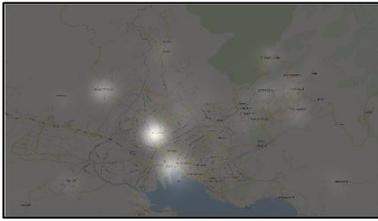
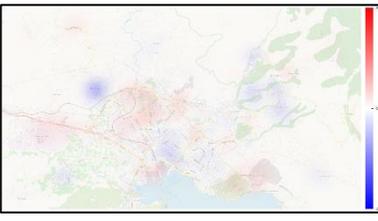
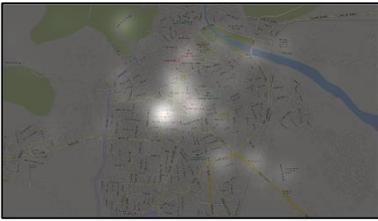
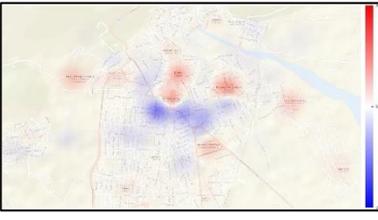
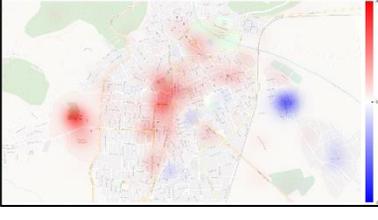
	Google - OSM			
	OSM - Wikimedia			
Larissa 12z	Bing - ESRI			
	ESRI - Google			
	Google - OSM			

	OSM - Wikimedia			
Patras 12z	Bing - ESRI			
	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			

Serres 14z	Bing - ESRI			
	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			
Thessaloniki 12z	Bing - ESRI			

	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			
Trikala 14z	Bing - ESRI			
	ESRI - Google			

Volos 14z	Google - OSM			
	OSM - Wikimedia			
	Bing - ESRI			
	ESRI - Google			
	Google - OSM			
	Google - OSM			

	OSM - Wikimedia			
Xanthi 16z	Bing - ESRI			
	ESRI - Google			
	Google - OSM			
	OSM - Wikimedia			

Chapter 6 : Conclusion and future work

6.1 Conclusion

The generated grayscale statistical heatmaps visualize the salient objects in each visual stimulus. More specifically, utilizing the EyeMMV toolbox, the visualization of binocular gaze data of 30 participants was achieved, revealing the locations that are more probable to be observed in 75 cartographic backgrounds from distinct online map services and distinct zoom levels.

The indices calculated presented a uniform result, meaning that the same cartographic design principles are followed to visualize the elements and the objects on visual stimuli and require the same amount of attention and cognitive load, in order to be explored by the participants. However, the grayscale statistical heatmaps demonstrate, that in fact the visual stimuli differ from each other, having distinct salient locations in each platform. The comparisons between grayscale statistical heatmaps of the same geographical area, but different service, revealed how the visual exploration happened and where the attention of the participants was drawn.

Visual behavior differs in each zoom level. Observing the grayscale statistical heatmaps, the exploration pattern seems to cover larger areas in cartographic backgrounds of higher zoom levels compared with visual scenes of smaller zoom level. The level of abstraction in the scale of 12z zoom level is much higher than the 16z zoom level. Having less details in 12z zoom level, the salient objects detected were the labels with large fonts, while in scales of 16z and 14z zoom levels, the design is more detailed and the participants detected more salient objects.

The comparison heatmaps revealed differences in each service depicting the same cartographic background. Although the visual scenes are in the same scale and the online map service depicts the same area, the salient objects located by the observers differ. The symbolization used is distinct in each platform, therefore more prominent objects were the ones with contrasting hue, different shape compared to the rest of the elements, or labels with larger font.

Evidently, the participants were searching for the label of each city's name, of which the visual scene was displayed. Although, in general, the visual search begun from the center, the pattern followed was guided by objects higher in the cartographic design hierarchy. Prominent objects, such as labels with large and bold fonts, worked as distractors for the observers. There wasn't a specific

task to be done, because the experiment was conducted in free viewing conditions, but if cartographic design principles were followed, in some cases, other objects should be salient or more prominent.

Ultimately, cartographic design can affect the allocation of the visual attention of the observers and determine the salient objects of a map. The color scheme used, the manipulation of the visual variables and the principles followed for a balanced cartographic representation, can create a hierarchical order in the elements of the map, which allows the user to observe the ones that are prominent, and afterwards all the “secondary” details. Eventually, what found to be a salient object by the objective observation of the participants, should be further investigated, for the visual stimuli should be examined if it is a pleasant and simultaneously an effective and efficient cartographic product.

6.2 Future perspectives

The grayscale statistical heatmaps produced, can serve as an objective ground truth, able to feed deep learning algorithms, in order to create user-oriented maps for every task and need. Moreover, further research can be made to understand how cartographic design on each online map service affects the visual perception. Analysis on the color schemes, the visual variables and the hierarchical order of the visualized objects of the scene, relatively to the observers’ gaze pattern. Furthermore, the indices calculated from the fixation and saccade report, can be statistically analyzed for more accurate results and gain insights just from the indices, without necessarily producing grayscale statistical heatmaps.

The grayscale statistical heatmap comparisons could be achieved in other ways. The combinations that can be achieved are $C(n,k) = n! / (k!(n-k)!)$ with n being the total number of grayscale statistical heatmaps and k is the number of heatmaps selected at a time. Given that we have five grayscale statistical heatmaps of a geographical area for distinct services, and the comparison is made by subtracting one heatmap from the other, there are ten different combinations, instead of four that have been made in this study. Hence, more combinations provide more information on the allocation of the visual attention on grayscale statistical heatmaps of distinct online map services, but on the same zoom level.

Of course, the search for salient objects can be achieved with a combination of methods. Along with eye movement tracking, mouse tracking can also be implemented, for an interactive search.

Chapter 6 – Conclusion and future work

The cartographic backgrounds could be observed in the web page of the platform, in real time, and not as an image, so usability of the platform can be tested simultaneously.

Chapter 7 : References

7.1 Bibliography

- Alaçam, Ö., & Dalc, M. (2009). A usability study of webmaps with eye tracking tool: The effects of iconic representation of information. *Human-Computer Interaction. New Trends: 13th International Conference, HCI International 2009, San Diego, CA, USA, July 19-24, 2009, Proceedings, Part I 13*, 12–21.
- Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., Jern, M., Kraak, M. J., Schumann, H., & Tominski, C. (2010). Space, time and visual analytics. *International Journal of Geographical Information Science*, 24(10), 1577–1600.
<https://doi.org/10.1080/13658816.2010.508043>
- Andrienko, G. L., & Andrienko, N. V. (1999). Interactive maps for visual data exploration. *International Journal of Geographical Information Science*, 13(4), 355–374.
- Arba-Mosquera, S., & Aslanides, I. M. (2012). Analysis of the effects of Eye-Tracker performance on the pulse positioning errors during refractive surgery. *Journal of Optometry*, 5(1), 31–37.
- Aserinsky, E., & Kleitman, N. (1953). Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. *Science*, 118(3062), 273–274.
- Bargiota, T., Mitropoulos, V., Krassanakis, V., & Nakos, B. (2013). Measuring locations of critical points along cartographic lines with eye movements. *Proceedings of the 26th International Cartographic Association Conference, Dresden, Germany*.
- Bartuzel, M. M., Wróbel, K., Tamborski, S., Meina, M., Nowakowski, M., Dalasiński, K., Szkulmowska, A., & Szkulmowski, M. (2020). High-resolution, ultrafast, wide-field retinal eye-tracking for enhanced quantification of fixational and saccadic motion. *Biomedical Optics Express*, 11(6), 3164–3180.
- Batista Narcizo, F., & Witzner Hansen, D. (2015). Depth compensation model for gaze estimation in sport analysis. *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 71–78.
- Bengi, H., & Thomas, J. G. (1968). Three electronic methods for recording ocular tremor. *Medical and Biological Engineering*, 6, 171–179.
- Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, 303(5919), 696–698.
- Bertin, J. (1983). *Semiology of graphics*. University of Wisconsin press.

- Boulos, M. N. K., & Burden, D. (2007). Web GIS in practice V: 3-D interactive and real-time mapping in Second Life. In *International Journal of Health Geographics* (Vol. 6, Issue 1, pp. 1–16). BioMed Central.
- Bour, L. J., Van Gisbergen, J. A. M., Bruijns, J., & Ottes, F. P. (1984). The double magnetic induction method for measuring eye movement-results in monkey and man. *IEEE Transactions on Biomedical Engineering*, 5, 419–427.
- Bowers, N. R., Boehm, A. E., & Roorda, A. (2019). The effects of fixational tremor on the retinal image. *Journal of Vision*, 19(11), 8.
- Braun, J., & Julesz, B. (1998). Withdrawing attention at little or no cost: detection and discrimination tasks. *Perception & Psychophysics*, 60(1), 1–23.
- Braun, J., & Sagi, D. (1990). Vision outside the focus of attention. *Perception & Psychophysics*, 48(1), 45–58.
- Brewer, C. A., & Buttenfield, B. P. (2007). Framing guidelines for multi-scale map design using databases at multiple resolutions. *Cartography and Geographic Information Science*, 34(1), 3–15. <https://doi.org/10.1559/152304007780279078>
- Brodersen, L. (2003). Modelling the visualization of internet maps. In *Maps and the Internet* (pp. 421–436). Elsevier.
- Brown, M., Marmor, M., Vaegan, Zrenner, E., Brigell, M., & Bach, M. (2006). ISCEV standard for clinical electro-oculography (EOG) 2006. *Documenta Ophthalmologica*, 113, 205–212.
- Brychtova, A., & Coltekin, A. (2016). An Empirical User Study for Measuring the Influence of Colour Distance and Font Size in Map Reading Using Eye Tracking. *Cartographic Journal*, 53(3), 202–212. <https://doi.org/10.1179/1743277414Y.0000000103>
- Brychtova, A., Popelka, S., & Dobesova, Z. (2012). Eye-tracking methods for investigation of cartographic principles. *SGEM 2012 Conference Proceedings*, 2, 1041–1048.
- Burch, M., Blascheck, T., Kurzhals, K., Pflüger, H., Raschke, M., Weiskopf, D., & Pfeiffer, T. (2015). Eye Tracking Visualization. *Eurographics (Tutorials)*.
- Buswell, G. T. (1935). *How people look at pictures: a study of the psychology and perception in art*.
- Buttenfield, B. P. (2002). Transmitting vector geospatial data across the Internet. *International Conference on Geographic Information Science*, 51–64.
- Cartwright, W. E. (2008). Mapping in a digital age. *The Handbook of Geographic Information Science*, 199–221.
- Cheng, M.-M., Mitra, N. J., Huang, X., Torr, P. H. S., & Hu, S.-M. (2014). Global contrast based salient region detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3), 569–582.

- Choe, K. W., Blake, R., & Lee, S.-H. (2016). Pupil size dynamics during fixation impact the accuracy and precision of video-based gaze estimation. *Vision Research*, *118*, 48–59.
- Ciołkosz-Styk, A. (2012). The visual search method in map perception research. *Geoinformation Issues*, *4*(1), 33–42.
- Clarke, K. C. (2004). Mobile mapping and geographic information systems. *Cartography and Geographic Information Science*, *31*(3), 131–136.
- Collewijn, H. (2001). Interocular timing differences in the horizontal components of human saccades. *Vision Research*, *41*(25–26), 3413–3423.
- Collewijn, H., der Mark, F., & Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research*.
- Collewijn, H., der Steen, J., Ferman, L., & Jansen, T. C. (1985). Human ocular counterroll: assessment of static and dynamic properties from electromagnetic scleral coil recordings. *Experimental Brain Research*, *59*, 185–196.
- Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1988). Binocular co-ordination of human vertical saccadic eye movements. *The Journal of Physiology*, *404*(1), 183–197.
- Çöltekin, A., Heil, B., Garlandini, S., & Fabrikant, S. I. (2009). Evaluating the effectiveness of interactive map interface designs: a case study integrating usability metrics with eye-movement analysis. *Cartography and Geographic Information Science*, *36*(1), 5–17.
- Cornsweet, T. N. (1958). New technique for the measurement of small eye movements. *JOSA*, *48*(11), 808–811.
- Crampton, J. W. (2009). Cartography: maps 2.0. *Progress in Human Geography*, *33*(1), 91–100.
- Crane, H. D., & Steele, C. M. (1985). Generation-V dual-Purkinje-image eyetracker. *Applied Optics*, *24*(4), 527–537.
- Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, *8*(2), 97–107.
- Cromley, R. G., & Campbell, G. M. (1992). Integrating quantitative and qualitative aspects of digital line simplification. *The Cartographic Journal*, *29*(1), 25–30.
- Cybulski, P., & Krassanakis, V. (2022). The effect of map label language on the visual search of cartographic point symbols. *Cartography and Geographic Information Science*, *49*(3), 189–204. <https://doi.org/10.1080/15230406.2021.2007419>
- Deubel, H., & Bridgeman, B. (1995). Fourth Purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades. *Vision Research*, *35*(4), 529–538.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837.

- DiBiase, D., MacEachren, A. M., Krygier, J. B., & Reeves, C. (1992). Animation and the role of map design in scientific visualization. *Cartography and Geographic Information Systems*, 19(4), 201–214.
- Dolezalova, J., & Popelka, S. (2016). The Use of Simple Graphs and Cliques for Analysis of Cartographic Eye-Tracking Data. *ITAT*, 206–211.
- Dong, W., Jiang, Y., Zheng, L., Liu, B., & Meng, L. (2018). Assessing map-reading skills using eye tracking and Bayesian structural equation modelling. *Sustainability (Switzerland)*, 10(9). <https://doi.org/10.3390/su10093050>
- Dong, W., Liao, H., Xu, F., Liu, Z., & Zhang, S. (2014). Using eye tracking to evaluate the usability of animated maps. *Science China Earth Sciences*, 57, 512–522.
- Droege, D., & Paulus, D. (2010). Pupil center detection in low resolution images. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 169–172.
- Duchowski, A. (2007). Eye Tracking Methodology: Theory and Practice. In *Eye Tracking Methodology: Theory and Practice*. <https://doi.org/10.1007/978-1-84628-609-4>
- Duwaer, A. L., Van Den Brink, G., Van Antwerpen, G., & Keemink, C. J. (1982). Comparison of subjective and objective measurements of ocular alignment in the vertical direction. *Vision Research*, 22(8), 983–989.
- Eckert, M. (1921). Die Kartographie im Kriege. *Geographische Zeitschrift*, 27(1./2. H), 18–28.
- Edler, D., Keil, J., Tuller, M.-C., Bestgen, A.-K., & Dickmann, F. (2020). Searching for the ‘right’ legend: The impact of legend position on legend decoding in a cartographic memory task. *The Cartographic Journal*, 57(1), 6–17.
- Elwood, S. (2010). Geographic information science: Emerging research on the societal implications of the geospatial web. *Progress in Human Geography*, 34(3), 349–357.
- Fabrikant, S. I., & Goldsberry, K. (2005). Thematic relevance and perceptual salience of dynamic geovisualization displays. *Proceedings, 22th ICA/ACI International Cartographic Conference, A Coruña, Spain*.
- Fabrikant, S. I., & Lobben, A. (2009). Introduction: Cognitive issues in geographic information visualization. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44(3), 139–143.
- Fabrikant, S. I., Rebich Hespahanha, S., & Hegarty, M. (2010). C 2010 by Association of American Geographers Initial submission. In *Annals of the Association of American Geographers* (Vol. 100, Issue 1).
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178–210.

- Friedmannová, L., Konečný, M., & Staněk, K. (2006). An adaptive cartographic visualization for support of the crisis management. *Proceedings of AutoCarto*, 100–105.
- Garlandini, S., & Fabrikant, S. I. (2009). Evaluating the effectiveness and efficiency of visual variables for geographic information visualization. *Spatial Information Theory: 9th International Conference, COSIT 2009 Aber Wrac'h, France, September 21-25, 2009 Proceedings 9*, 195–211.
- Godara, P., Dubis, A. M., Roorda, A., Duncan, J. L., & Carroll, J. (2010). Adaptive optics retinal imaging: emerging clinical applications. *Optometry and Vision Science: Official Publication of the American Academy of Optometry*, 87(12), 930.
- Gökstorp, S. G. E., & Breckon, T. P. (2022). Temporal and non-temporal contextual saliency analysis for generalized wide-area search within unmanned aerial vehicle (UAV) video. *The Visual Computer*, 38(6), 2033–2040.
- Goldberg, J. H., & Wichansky, A. M. (2003). Chapter 23 - Eye Tracking in Usability Evaluation: A Practitioner's Guide. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The Mind's Eye* (pp. 493–516). North-Holland. <https://doi.org/https://doi.org/10.1016/B978-044451020-4/50027-X>
- Goldsberry, K. P. (2007). *Real-time traffic maps*. University of California, Santa Barbara.
- Golebiowska, I., Opach, T., & Rød, J. K. (2020). Breaking the Eyes: How Do Users Get Started with a Coordinated and Multiple View Geovisualization Tool? *The Cartographic Journal*, 57(3), 235–248.
- Goodchild, M. F. (2007). Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. *International Journal of Spatial Data Infrastructures Research*, 2(2), 24–32.
- Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, 14(5), 462–466.
- Guestrin, E. D., & Eizenman, M. (2006). General theory of remote gaze estimation using the pupil center and corneal reflections. *IEEE Transactions on Biomedical Engineering*, 53(6), 1124–1133.
- Haklay, M., Singleton, A., & Parker, C. (2008). Web mapping 2.0: The neogeography of the GeoWeb. *Geography Compass*, 2(6), 2011–2039. <https://doi.org/10.1111/j.1749-8198.2008.00167.x>
- Hansen, D. W., & Ji, Q. (2009). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3), 478–500.
- Hansen, D. W., & Pece, A. E. C. (2005). Eye tracking in the wild. *Computer Vision and Image Understanding*, 98(1), 155–181.
- Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. *Advances in Neural Information Processing Systems*, 19.

- Harvey, F. (2012). To volunteer or to contribute locational information? Towards truth in labeling for crowdsourced geographic information. In *Crowdsourcing geographic knowledge: Volunteered geographic information (VGI) in theory and practice* (pp. 31–42). Springer.
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7(11), 498–504.
- Henderson, J. M., Brockmole, J. R., Castelano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In *Eye movements* (pp. 537–III). Elsevier.
- Henderson, J. M., & Ferreira, F. (2004). *Scene perception for psycholinguists*.
- Holmqvist, K. (2016). Optimal settings for commercial event detection algorithms based on the level of noise. In *Behavior Research Methods*, 1–21.
- Holmqvist, K., & Andersson, R. (2017). Eye tracking: A comprehensive guide to methods. *Paradigms and Measures*.
- Holmqvist, K., & Blignaut, P. (2020). Small eye movements cannot be reliably measured by video-based P-CR eye-trackers. *Behavior Research Methods*, 52, 2098–2121.
- Holmqvist, K., Örbom, S. L., Hooge, I. T. C., Niehorster, D. C., Alexander, R. G., Andersson, R., Benjamins, J. S., Blignaut, P., Brouwer, A. M., Chuang, L. L., Dalrymple, K. A., Drieghe, D., Dunn, M. J., Ettinger, U., Fiedler, S., Foulsham, T., van der Geest, J. N., Hansen, D. W., Hutton, S. B., ... Hessels, R. S. (2023). Eye tracking: empirical foundations for a minimal reporting guideline. *Behavior Research Methods*, 55(1), 364–416. <https://doi.org/10.3758/s13428-021-01762-8>
- Hooge, I., Holmqvist, K., & Nyström, M. (2016). The pupil is faster than the corneal reflection (CR): Are video based pupil-CR eye trackers suitable for studying detailed dynamics of eye movements? *Vision Research*, 128, 6–18.
- Horbiński, T. (2019). Progressive evolution of designing internet maps on the example of Google Maps. *Geodesy and Cartography*, 68(1).
- Houben, M. M. J., Goumans, J., & van der Steen, J. (2006). Recording three-dimensional eye movements: scleral search coils versus video oculography. *Investigative Ophthalmology & Visual Science*, 47(1), 179–187.
- Incoul, A., Ooms, K., & De Maeyer, P. (2015). Comparing paper and digital topographic maps using eye tracking. *Modern Trends in Cartography: Selected Papers of CARTOCON 2014*, 339–356.
- Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. *Eye Guidance in Reading and Scene Perception*, 29–53.
- Intoy, J., & Rucci, M. (2020). Finely tuned eye movements enhance visual acuity. *Nature Communications*, 11(1), 795.

- Irving, E. L., Zacher, J. E., Allison, R. S., & Callender, M. G. (2003). Effects of scleral search coil wear on visual function. *Investigative Ophthalmology & Visual Science*, 44(5), 1933–1938.
- Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*, 217, 105–133.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. In *Vision Research* (Vol. 40). www.elsevier.com
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254–1259.
- Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The mind's eye* (pp. 573–605). Elsevier.
- Jaschinski, W. (2016). Pupil size affects measures of eye position in video eye tracking: implications for recording vergence accuracy. *Journal of Eye Movement Research*, 9(4).
- Jaschinski, W., Jainta, S., & Kloke, W. B. (2010). Objective vs subjective measures of fixation disparity for short and long fixation periods. *Ophthalmic and Physiological Optics*, 30(4), 379–390.
- Jones, C. B., Abdelmoty, A. I., Lonergan, M. E., van der Poorten, P., & Zhou, S. (2000). Multi-scale spatial database design for online generalisation. *9th International Symposium on Spatial Data Handling*, 34–44.
- Keates, J. S. (2014). *Understanding maps*. Routledge.
- Keskin, M., Krassanakis, V., & Çöltekin, A. (2023). Visual Attention and Recognition Differences Based on Expertise in a Map Reading and Memorability Study. *ISPRS International Journal of Geo-Information*, 12(1). <https://doi.org/10.3390/ijgi12010021>
- Keskin, M., Ooms, K., Dogru, A. O., & De Maeyer, P. (2020). Exploring the cognitive load of expert and novice map users using EEG and eye tracking. *ISPRS International Journal of Geo-Information*, 9(7), 429.
- Kiefer, P., Giannopoulos, I., Raubal, M., & Duchowski, A. (2017). Eye tracking for spatial research: Cognition, computation, challenges. *Spatial Cognition & Computation*, 17(1–2), 1–19.
- Ko, H., Snodderly, D. M., & Poletti, M. (2016). Eye movements between saccades: Measuring ocular drift and tremor. *Vision Research*, 122, 93–104.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219–227.
- Krassanakis, V. (2009). *Recording the trace of visual search: a research method of the selectivity of hole as a basic shape characteristic (in Greek)*. <https://doi.org/http://dx.doi.org/10.26240/heal.ntua.4447>
- Krassanakis, V. (2021). Aggregated Gaze Data Visualization Using Contiguous Irregular Cartograms. *Digital*, 1(3), 130–144. <https://doi.org/10.3390/digital1030010>

- Krassanakis, V., & Cybulski, P. (2019). A review on eye movement analysis in map reading process: the status of the last decade. *Geodesy and Cartography*, 68, 191–209. <https://doi.org/10.24425/gac.2019.126088>
- Krassanakis, V., & Cybulski, P. (2021). Eye tracking research in cartography: Looking into the future. *ISPRS International Journal of Geo-Information*, 10(6). <https://doi.org/10.3390/ijgi10060411>
- Krassanakis, V., & Filippakopoulou, V. (2023). *Spatial representations: Creation techniques, theoretical principles, evaluation methods (in Greek)* (K. Pitsoli & L.-M. Misthos, Eds.). Kallipos, Open Academic Editions. <https://doi.org/http://dx.doi.org/10.57713/kallipos-160>
- Krassanakis, V., Filippakopoulou, V., & Nakos, B. (2014). EyeMMV toolbox: An eye movement post-analysis tool based on a two-step spatial dispersion threshold for fixation identification. *Journal of Eye Movement Research*, 7(1). <https://doi.org/10.16910/jemr.7.1.1>
- Krassanakis, V., Lelli, A., Lokka, I., Filippakopoulou, V., & Nakos, B. (2013, July). *Searching for salient locations in topographic maps*. https://doi.org/10.2390/biecoll-saga2013_0
- Krygier, J. B. (1994). Sound and geographic visualization. In *Modern cartography series* (Vol. 2, pp. 149–166). Elsevier.
- Laube, P., Purves, R. S., & Haklay, M. (2010). *Cross-scale movement trajectory analysis*.
- Leclerc, Y. G., Reddy, M., Iverson, L., & Eriksen, M. (2001). The geoweb—a new paradigm for finding data on the web. *Proceedings of the International Cartographic Conference*.
- Liao, Y., Wang, J., Jaehning, E. J., Shi, Z., & Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Research*, 47(W1), W199–W205.
- Liu, B., Dong, W., Zhan, Z., Wang, S., & Meng, L. (2020). Differences in the gaze behaviours of pedestrians navigating between regular and irregular road patterns. *ISPRS International Journal of Geo-Information*, 9(1). <https://doi.org/10.3390/ijgi9010045>
- Liversedge, S. P., White, S. J., Findlay, J. M., & Rayner, K. (2006). Binocular coordination of eye movements during reading. *Vision Research*, 46(15), 2363–2374.
- MacEachren, A. M. (2004). *How maps work: representation, visualization, and design*. Guilford Press.
- Mackaness, W. A., & Ruas, A. (2007). Evaluation in the map generalisation process. In *Generalisation of geographic information* (pp. 89–111). Elsevier.
- Malpeli, J. G. (1998). Measuring eye position with the double magnetic induction method. *Journal of Neuroscience Methods*, 86(1), 55–61.
- Manson, S. M., Kne, L., Dyke, K. R., Shannon, J., & Eria, S. (2012). Using eye-tracking and mouse metrics to test usability of web mapping navigation. *Cartography and Geographic Information Science*, 39(1), 48–60.

- Mardanbegi, D., Hansen, D. W., & Pederson, T. (2012). Eye-based head gestures. *Proceedings of the Symposium on Eye Tracking Research and Applications*, 139–146.
- McCamy, M. B., Collins, N., Otero-Millan, J., Al-Kalbani, M., Macknik, S. L., Coakley, D., Troncoso, X. G., Boyle, G., Narayanan, V., Wolf, T. R., & others. (2013). Simultaneous recordings of ocular microtremor and microsaccades with a piezoelectric sensor and a video-oculography system. *PeerJ*, 1, e14.
- Meng, L., & Reichenbacher, T. (2005). Map-based mobile services. In *Map-based mobile services: Theories, methods and implementations* (pp. 1–10). Springer.
- Montello, D. R. (2002). Cognitive map-design research in the twentieth century: Theoretical and empirical approaches. *Cartography and Geographic Information Science*, 29(3), 283–304.
- Mulligan, J. B. (1997). Image processing for improved eye-tracking accuracy. *Behavior Research Methods, Instruments, & Computers*, 29(1), 54–65.
- Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, 29(11), 1631–1647.
- Nasiopoulos, E., Risko, E. F., & Kingstone, A. (2015). Social attention, social presence, and the dual function of gaze. *The Many Faces of Social Attention: Behavioral and Neural Measures*, 129–155.
- Niehorster, D.C., Zemblys, R., Beelders, T. *et al.* Characterizing gaze position signals and synthesizing noise during fixations in eye-tracking data. *Behav Res* 52, 2515–2534 (2020). <https://doi.org/10.3758/s13428-020-01400-9>
- Nuthmann, A., & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision*, 9(5), 31.
- Nyström, M., Andersson, R., Holmqvist, K. *et al.* The influence of calibration method and eye physiology on eyetracking data quality. *Behav Res* 45, 272–288 (2013). <https://doi.org/10.3758/s13428-012-0247-4>
- Nyström, M., Hooge, I., & Andersson, R. (2016). Pupil size influences the eye-tracker signal during saccades. *Vision Research*, 121, 95–103.
- Oliva, M., Niehorster, D. C., Jarodzka, H., & Holmqvist, K. (2017). Influence of coactors on saccadic and manual responses. *I-Perception*, 8(1), 2041669517692814.
- Olshausen, B. A., Anderson, C. H., & Van Essen, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, 13(11), 4700–4719.
- Opach, T., & Nossum, A. (2011). Evaluating the usability of cartographic animations with eye-movement analysis. *25th International Cartographic Conference*, 11.

- O'regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of 'mudsplashes.' *Nature*, *398*(6722), 34.
- Perrin, A.-F.; Krassanakis, V.; Zhang, L.; Ricordel, V.; Perreira Da Silva, M.; Le Meur, O. EyeTrackUAV2: A Large-Scale Binocular Eye-Tracking Dataset for UAV Videos. *Drones* **2020**, *4*, 2. <https://doi.org/10.3390/drones4010002>
- Peterson, M. P. (2008). Trends in internet and ubiquitous cartography. *Cartographic Perspectives*, *61*, 36–49.
- Poole, A., & Ball, L. J. (2005). Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. *The Encyclopedia of Human Computer Interaction*, 211–219.
- Poole, A., Ball, L. J., & Phillips, P. (2005). In search of salience: A response-time and eye-movement analysis of bookmark recognition. *People and Computers XVIII—Design for Life: Proceedings of HCI 2004*, 363–378.
- Popelka, S., Vondrakova, A., & Hujnakova, P. (2019). Eye-tracking Evaluation of Weather Web Maps. *ISPRS International Journal of Geo-Information*, *8*(6). <https://doi.org/10.3390/ijgi8060256>
- Putnam, N. M., Hofer, H. J., Doble, N., Chen, L., Carroll, J., & Williams, D. R. (2005). The locus of fixation and the foveal cone mosaic. *Journal of Vision*, *5*(7), 3.
- Räihä, K.-J., Aula, A., Majaranta, P., Rantala, H., & Koivunen, K. (2005). Static visualization of temporal eye-tracking data. *Human-Computer Interaction-INTERACT 2005: IFIP TC13 International Conference, Rome, Italy, September 12-16, 2005. Proceedings 10*, 946–949.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372.
- Reichenbacher, T. (2003). Adaptive methods for mobile cartography. *Proceedings of the 21st International Cartographic Conference*, 1311–1322.
- Reulen, J. P. H., & Bakker, L. (1982). The measurement of eye movement using double magnetic induction. *IEEE Transactions on Biomedical Engineering*, *11*, 740–744.
- Robinson, A. H. (1995). *Elements of Cartography*. Wiley.
<https://books.google.com.na/books?id=vu9OAAAAMAAJ>
- Robinson, A. H., & Petchenik, B. B. (1976). *The nature of maps: Essays toward understanding maps and mapping*. University of Chicago Press Chicago.
- Robinson, D. A. (1963). A method of measuring eye movement using a scieral search coil in a magnetic field. *IEEE Transactions on Bio-Medical Electronics*, *10*(4), 137–145.

- Romero-Borja, F., Venkateswaran, K., Roorda, A., & Hebert, T. (2005). Optical slicing of human retinal tissue in vivo with the adaptive optics scanning laser ophthalmoscope. *Applied Optics*, 44(19), 4032–4040.
- Roth, R. E. (2013). An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), 2356–2365.
- Roth, R. E., Brewer, C. A., & Stryker, M. S. (2011). A typology of operators for maintaining legible map designs at multiple scales. *Cartographic Perspectives*, 68, 29–64.
- Roth, R. E., Donohue, R. G., Sack, C. M., Wallace, T. R., & Buckingham, T. M. A. (2014). A process for keeping pace with evolving web mapping technologies. *Cartographic Perspectives*, 2014(78), 25–52. <https://doi.org/10.14714/CP78.1273>
- Roth, R. E., Robinson, A., Stryker, M., MacEachren, A. M., Lengerich, E. J., & Koua, E. (2008). Web-based geovisualization and geocollaboration: Applications to public health. *2008 Joint Statistical Meeting, Invited Session on Web Mapping*, 2–5.
- Sack, C. M. (2013). *Mapmaking for Change: Online Participatory Mapping Tools for Revealing Landscape Values in the Bad River Watershed*.
- Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, 1(7), 261–267.
- Skopeliti, A., & Stamou, L. (2019). Online map services: Contemporary cartography or a new cartographic culture? *ISPRS International Journal of Geo-Information*, 8(5). <https://doi.org/10.3390/ijgi8050215>
- SR Research Ltd. (2020). *SR Research Experiment Builder 2.3.1*. <https://www.sr-support.com/forums/showthread.php?t=99>.
- SR Research Ltd. (2021a). *EyeLink® 1000 Plus User Manual Desktop, LCD Arm, Tower, Primate and Long Range Mounts Remote, 2000 Hz and Fiber Optic Camera Upgrades*. <http://www.sr-research.com>
- SR Research Ltd. (2021b). *EyeLink® Data Viewer 4.2.1*.
- Steinke, T. R. (1987). Eye movement studies in cartography and related fields. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 24(2), 40–73.
- Stevenson, S. B., Sheehy, C. K., & Roorda, A. (2016). Binocular eye tracking with the tracking scanning laser ophthalmoscope. *Vision Research*, 118, 98–104.
- Švede, A., Treija, E., Jaschinski, W., & Krūmiņa, G. (2015). Monocular versus binocular calibrations in evaluating fixation disparity with a video-based eye-tracker. *Perception*, 44(8–9), 1110–1128.
- Swienty, O., Kurz, F., & Reichenbacher, T. (2007). Attention Guiding Visualisation in Remote Sensing IIM Systems. *Photogrammetrie Fernerkundung Geoinformation*, 4, 239–252.

- Tatler, B. W., Hansen, D. W., & Pelz, J. B. (2019). Eye movement recordings in natural settings. *Eye Movement Research: An Introduction to Its Scientific Foundations and Applications*, 549–592.
- Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review*, 14, 663–668.
- Toet, A., Kooi, F. L., Bijl, P., & Valeton, J. M. (1998). Visual conspicuity determines human target acquisition performance. *Optical Engineering*, 37(7), 1969–1975.
- Tong, N., Lu, H., Zhang, L., & Ruan, X. (2014). Saliency detection with multi-scale superpixels. *IEEE Signal Processing Letters*, 21(9), 1035–1039.
- Torok, N., Guillemin Jr, V., & Barnothy, J. M. (1951). Lxxx photoelectric nystagmography. *Annals of Otolaryngology, Rhinology & Laryngology*, 60(4), 917–926.
- Torralba, A., Oliva, A., Castelhana, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological Review*, 113(4), 766.
- Treisman, A. (1998). The perception of features and objects. *Visual Attention*, 8, 26–54.
- Ullah, R., Mengistu, E. Z., Van Elzakker, C., & Kraak, M.-J. (2016). Usability evaluation of centered time cartograms. *Open Geosciences*, 8(1), 337–359.
- Wade, N., & Tatler, B. W. (2005). *The moving tablet of the eye: The origins of modern eye movement research*. Oxford University Press.
- Wang, X., Holmqvist, K., & Alexa, M. (2019). The mean point of vergence is biased under projection. *Journal of Eye Movement Research*, 12(4).
- Wang, Y., Mehler, B., Reimer, B., Lammers, V., D'Ambrosio, L. A., & Coughlin, J. F. (2010). The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics*, 53(3), 404–420.
- Wilson, M. W. (2012). Location-based services, conspicuous mobility, and the location-aware future. *Geoforum*, 43(6), 1266–1275.
- Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202–238.
- Wood, J. (2005). Multim im parvo—many things in a small place. *Exploring Geovisualization*, 313–324.
- Woodward, D., & Lewis, G. M. (1998). *Cartography in the Traditional African! J Aterican! J Arctic! J Australian! J and Pacific Societies*.
- Wyatt, H. J. (2010). The human pupil and the use of video-based eyetrackers. *Vision Research*, 50(19), 1982–1988.

Yarbus, A. L., & Yarbus, A. L. (1967). Eye movements during perception of complex objects. *Eye Movements and Vision*, 171–211.

7.2 Web sources

ET4S. (2023). Retrieved from https://et4s.ethz.ch/2020/?page_id=53 (last visited: 17/07/2023)

Hellenic Statistical Authority. (2023). Retrieved from <https://www.statistics.gr/en/home/> (last visited: 17/07/2023)

ICA (International Cartographic Association). (2012, 12 5). Retrieved from Invitation to workshop “Eye tracking: why, when, and how?”: <https://icaci.org/invitation-to-workshop-eye-tracking-why-when-and-how/> (last visited: 17/07/2023)

ICA (International Cartographic Association). (2020, 12 1). Retrieved from <https://icaci.org/strategic-plan/>

ICA Commission on Cognitive Issues in Geographic Information Visualization. (2019, 12 10). Retrieved from <https://cogvis.icaci.org/> (last visited: 17/07/2023)

Appendix A

A.1 Fixation duration indices

```

clc
clear
format long g
%-----%
%-----%
%% LOAD DATA

% Fixation data
fixations = readtable("fixations.csv");

%% CALCULATE FIXATION DURATION FOR EVERY TRIAL BASED ON THE IMAGE

% Cell to string, the first column
fixations.RECORDING_SESSION_LABEL = string(fixations.RECORDING_SESSION_LABEL);

% Cell to table, images column
images = fixations.image;
images = cell2table(images);

% Get unique images for all trials
uniqueImages = unique(images, "rows");

% Number of images
N_Stim = numel(uniqueImages);

% Table to string, images column
uniqueImages.images = string(uniqueImages.images);

% Preallocate string array for better performance
fix_dur_all = strings(N_Stim,1);

% Loop through each image
for i = 1:N_Stim
    img = uniqueImages.images(i);
    for j = 1: length(fixations.CURRENT_FIX_DURATION)
        pat_img = contains(fixations.image, img); % Logical indexing
        img_fix_dur = fixations(pat_img == 1, :);
    end
    min_fix_dur_img = min(img_fix_dur.CURRENT_FIX_DURATION); % min fixation duration
    max_fix_dur_img = max(img_fix_dur.CURRENT_FIX_DURATION); % max fixation duration
    mean_fix_dur_img = mean(img_fix_dur.CURRENT_FIX_DURATION); % average fixation
duration
    median_fix_dur_img = median(img_fix_dur.CURRENT_FIX_DURATION); % median fixation
duration
    std_fix_dur_img = std(img_fix_dur.CURRENT_FIX_DURATION); % std fixation duration
    img_fix_dur_ind = [min_fix_dur_img max_fix_dur_img mean_fix_dur_img
median_fix_dur_img std_fix_dur_img]; % Export in a table

```

```

    img_fix_dur_ind_str = strcat(img, ",", num2str(img_fix_dur_ind(1)), ",",
num2str(img_fix_dur_ind(2)), ",", num2str(img_fix_dur_ind(3)), ",",
num2str(img_fix_dur_ind(4)), ",", num2str(img_fix_dur_ind(5)) );
    fix_dur_all(i) = img_fix_dur_ind_str;
    disp(i)
end

%[fix_dur_all; img_fix_dur_ind_str];

```

A.2 Total number of fixations indices

```

clc
clear
format long g
%-----%
%-----%
%% LOAD DATA

% Fixation data
fixations = readtable("fixations.csv");

%% CALCULATE NUMBER OF FIXATIONS FOR EVERY TRIAL BASED ON THE IMAGE

% Extract from fixation report total number of fixations
fix_num = fixations(:, ["RECORDING_SESSION_LABEL", "TRIAL_INDEX", "image",
"TRIAL_FIXATION_TOTAL"]);

% Cell to string, the first column
fix_num.RECORDING_SESSION_LABEL = string(fix_num.RECORDING_SESSION_LABEL);

% Get unique number of fixations for every trial
uniqueFix_num = unique(fix_num, "rows");

% Get unique images for all trials
images = uniqueFix_num.image;

% Cell to table, images column
images = cell2table(images);
uniqueImages = unique(images, "rows");

% Number of images
N_Stim = numel(uniqueImages);

% Table to string, images column
uniqueImages.images = string(uniqueImages.images);

% Preallocate string array for better performance
fix_num_all = strings(N_Stim,1);

% Loop through each image
for i = 1:N_Stim
    img = uniqueImages.images(i);
    pat_img = contains(uniqueFix_num.image, img); % Logical indexing
    img_tot_num = uniqueFix_num(pat_img == 1, :);

```

```

    min_fix_num_img = min(img_tot_num.TRIAL_FIXATION_TOTAL); % min num of fixations
    max_fix_num_img = max(img_tot_num.TRIAL_FIXATION_TOTAL); % max num of fixations
    mean_fix_num_img = mean(img_tot_num.TRIAL_FIXATION_TOTAL); % average num of
fixations
    median_fix_num_img = median(img_tot_num.TRIAL_FIXATION_TOTAL); % median num of
fixations
    std_fix_num_img = std(img_tot_num.TRIAL_FIXATION_TOTAL); % std num of fixations
    img_fix_num_ind = [min_fix_num_img max_fix_num_img mean_fix_num_img
median_fix_num_img std_fix_num_img]; % Export in a table
    img_fix_num_ind_str = strcat(img, ",", num2str(img_fix_num_ind(1)), ",",
num2str(img_fix_num_ind(2)), ",", num2str(img_fix_num_ind(3)), ",",
num2str(img_fix_num_ind(4)), ",", num2str(img_fix_num_ind(5)) );
    fix_num_all(i) = img_fix_num_ind_str;
    disp(i)
end

```

A.3 Saccade amplitude indices

```

clc
clear
format long g
%------%
%------%
%% LOAD DATA

% Saccade data
saccades = readtable("saccades.csv");

%% CALCULATE SACCADE AMPLITUDE FOR EVERY TRIAL BASED ON THE IMAGE

% Cell to string, the first column
saccades.RECORDING_SESSION_LABEL = string(saccades.RECORDING_SESSION_LABEL);

% Cell to table, images column
images = saccades.image;
images = cell2table(images);

% Get unique images for all trials
uniqueImages = unique(images, "rows");

% Number of images
N_Stim = numel(uniqueImages);

% Table to string, images column
uniqueImages.images = string(uniqueImages.images);

% Preallocate string array for better performance
sac_amp_all = strings(N_Stim,1);

% Loop through each image
for i = 1:N_Stim
    img = uniqueImages.images(i);

```

```

    for j = 1:
length(saccades.CURRENT_SAC_AMPLITUDE(~isnan(saccades.CURRENT_SAC_AMPLITUDE))) % Loop
through each saccade amplitude value
        pat_img = contains(saccades.image, img); % Logical indexing
        img_sac_amp = saccades(pat_img == 1, :);
    end
        min_sac_amp_img =
min(img_sac_amp.CURRENT_SAC_AMPLITUDE(~isnan(img_sac_amp.CURRENT_SAC_AMPLITUDE))); %
min saccade amplitude
        max_sac_amp_img =
max(img_sac_amp.CURRENT_SAC_AMPLITUDE(~isnan(img_sac_amp.CURRENT_SAC_AMPLITUDE))); %
max saccade amplitude
        mean_sac_amp_img =
mean(img_sac_amp.CURRENT_SAC_AMPLITUDE(~isnan(img_sac_amp.CURRENT_SAC_AMPLITUDE))); %
average saccade amplitude
        median_sac_amp_img =
median(img_sac_amp.CURRENT_SAC_AMPLITUDE(~isnan(img_sac_amp.CURRENT_SAC_AMPLITUDE)));
% median saccade amplitude
        std_sac_amp_img =
std(img_sac_amp.CURRENT_SAC_AMPLITUDE(~isnan(img_sac_amp.CURRENT_SAC_AMPLITUDE))); %
std saccade amplitude
        img_sac_amp_ind = [min_sac_amp_img max_sac_amp_img mean_sac_amp_img
median_sac_amp_img std_sac_amp_img]; % Export in a table
        img_sac_amp_ind_str = strcat(img, ",", num2str(img_sac_amp_ind (1)), ",",
num2str(img_sac_amp_ind(2)), ",", num2str(img_sac_amp_ind(3)), ",",
num2str(img_sac_amp_ind(4)), ",", num2str(img_sac_amp_ind(5)) );
        sac_amp_all(i) = img_sac_amp_ind_str;
        disp(i)
    end
end

```

Appendix B

B.1 Heatmap generator

```

clc
clear
tic
% Directory with .txt files
directory = 'D:\SXOLI\Diploma - Thesis\output_cumulative';

% Folder to store heatmaps
outputFolder = 'D:\SXOLI\Diploma - Thesis\heatmaps_EyeMMV';
if ~exist(outputFolder, 'dir')
    mkdir(outputFolder);
end

% List of .txt files
fileList = dir(fullfile(directory, '*.txt'));

% Parameters for heatmap
spacing = 1; % Spacing between heatmap cells
max_hor = 1920; % Maximum horizontal value for heatmap
max_ver = 1080; % Maximum vertical value for heatmap
s_gaussian = 38; % Standard deviation for the Gaussian kernel
kernel_size_gaussian = 6 * s_gaussian; % Size of the Gaussian kernel for smoothing
the heatmap

% Loop for every .txt file
for i = 1:length(fileList)
    filename = fullfile(directory, fileList(i).name);

    % Data from the file
    data = readtable(filename);

    % Extract x and y coordinates
    x = data.Var7;
    y = data.Var8;

    % Coordinates as records
    records = [x, y];

    % Heatmap name
    [~, name, ~] = fileparts(filename);
    output_heatmap_name = fullfile(outputFolder, [name '_heatmap.png']);

    % Call function
    heatmap_generator_EyeMMV_modified_v3(records, output_heatmap_name, spacing,
max_hor, max_ver, kernel_size_gaussian, s_gaussian);
    fprintf('Heatmap Visualization is completed successfully\n')
end
toc

```

B.2 Heatmaps overlaid on visual stimuli

```

clc
clear

% Paths to the folders with the images
folder1 = 'D:\SXOLI\Diploma - Thesis\cartographic_backgrounds';
folder2 = 'D:\SXOLI\Diploma - Thesis\heatmaps_EyeMMV';

% List of images in folders
files1 = dir(fullfile(folder1, '*.png'));
numImg = numel(files1);

files2 = dir(fullfile(folder2, '*.png'));

% Set the desired output size
outputSize = [1920, 1080];

% Overlay images
for i = 1:numImg
    % Images from both folders
    imgName1 = fullfile(folder1, files1(i).name);
    img1 = imread(imgName1);

    % Extract the file name of the first image without the extension
    [~, Name, ~] = fileparts(imgName1);

    imgName2 = fullfile(folder2, files2(i).name);
    img2 = imread(imgName2);

    % Display first image
    imshow(img1);
    hold on;

    % Display second image and add transparency
    h = imshow(img2);
    set(h, 'AlphaData', 0.6);

    % Resize the overlaid image
    % outputImage = imresize(img2, outputSize);
    %

    % Save overlaid images
    outputFileName = fullfile('D:\SXOLI\Diploma - Thesis\heatmap_overlay', [Name,
    '_overlay.png']);
    % saveas(gcf, outputFileName);
    exportgraphics(gcf, outputFileName);
    hold off;
end

```

Appendix C

C.1 Heatmap comparison

```
clc
clear
close

% Directory paths
baseMapDir = 'D:\SXOLI\Diploma - Thesis\cartographic_backgrounds';
heatmapDir = 'D:\SXOLI\Diploma - Thesis\heatmaps_EyeMMV';

% Get the list of base map files
baseMapFiles = dir(fullfile(baseMapDir, '*.png'));

% Get the list of heatmap files
heatmapFiles = dir(fullfile(heatmapDir, '*.png'));

% Iterate over heatmaps
for j = 1:numel(heatmapFiles)

    if mod(j, 5) == 0
        continue;
    end

    % Read basemap
    base_map = imread(fullfile(baseMapDir, baseMapFiles(j).name));

    % Read heatmap
    heatmap1 = double(imread(fullfile(heatmapDir, heatmapFiles(j).name)));

    % Read heatmap2
    heatmap2 = double(imread(fullfile(heatmapDir, heatmapFiles(j+1).name)));

    % Calculate heatmap difference
    heatmap_dif = heatmap1 - heatmap2;

    % Create a figure
    plot = figure;

    imshow(base_map);
    hold on;

    % Display the difference heatmap with transparency
    h = imshow(heatmap_dif);
    set(h, 'AlphaData', 0.6);

    % Adjust plot properties
    hold off;
    axis off
    axis equal
```

```

%% Colormap of colorbar
% Define the colormap
numColors = 512; %(256+256)
colorMap = zeros(numColors, 3);

% Define the RGB values for the color gradient from blue to red
whiteColor = [1, 1, 1]; % White
blueColor = [0, 0, 1]; % Blue
redColor = [1, 0, 0]; % Red

% Calculate the RGB values for the color gradient
for i = 1:numColors/2
    colorMap(i, :) = blueColor + (i-1) * (whiteColor - blueColor) / (numColors/2-
1);
end

for i = numColors/2+1:numColors
    colorMap(i, :) = whiteColor + (i - numColors/2 - 1) * (redColor - whiteColor)
/ (numColors/2-1);
end

% Create a custom colorbar
colorbar;

% Apply the custom colormap to the colorbar
colormap(colorMap);

% Specify the tick locations and labels for the colorbar
caxis([-255, 255]);
cTicks = [-255, 0, 255];
cTickLabels = {'-255', '0', '255'};
cbar = colorbar;
cbar.Ticks = cTicks;
cbar.TickLabels = cTickLabels;

files = dir(fullfile(heatmapDir, '*.png'));
imgName = fullfile(heatmapDir, files(j).name);
[~, Name, ~] = fileparts(imgName);
outputFileName = fullfile('D:\SXOLI\Diploma - Thesis\heatmap_comparisons', [Name,
'_difference.png']);
% Save the figure or perform any desired operations
exportgraphics(plot, outputFileName)
fprintf(" %g , %g \n", j, j+1)

```