Spatio-temporal integration properties of the human visual system

Grillini, Alessandro

DOI:
10.33612/diss.136424282

Publication date:
2020

Chapter 8

General Discussion

In this thesis, I investigated the spatio-temporal integration properties of the human visual system using three experimental approaches: psychophysics, eye-tracking and computational modeling. In the following sections, I summarize the experimental findings and discuss possible interpretations and future medical application.

8.1 Summary of the findings

8.1.1 Spatio-temporal properties of eye-movements: description of algorithms and features

Chapter 2 introduced the methods and algorithms to extract the spatio-temporal features from eye-movement data and described their most relevant properties. The method, used in Chapters 3, 4, 5 and 6, consists of two target-tracking conditions (smooth and saccadic pursuits) to elicit the observers' ocular following response and visually-driven saccades. The resulting time-series of gaze positions were analyzed using the Eye-Movement Crosscorrelogram[2] and its extensions[26,79], which provide 10 features to quantify oculomotor performance from both a spatial and temporal perspective. For each pursuit modality (smooth pursuit and saccadic pursuit) and each cardinal axis, a set of spatio-temporal features is computed. In its most basic form (single contrast level, two pursuit modalities, horizontal and vertical axis analyzed separately) the analysis can provide 40 spatio-temporal features. At least in those without a disease affecting foveal vision, these features are not correlated with other "static" measures of visual functions such as visual acuity and contrast sensitivity.

8.1.2 Using the spatio-temporal properties of eye movements to classify visual field defects

In Chapter 3, we tested the hypothesis that there is a direct relationship between the presence of a visual field loss and the spatio-temporal properties of eye movements (computed as described in Chapter 2). We found, in healthy subjects using simulated scotoma, that each tested scotoma type (central loss, peripheral loss or hemifield loss) led to a characteristic profile of spatio-temporal features. Using machine learning, it was possible to accurately infer (91% accuracy) which type of scotoma caused certain patterns in the spatio-temporal features. Furthermore, the pattern found in the simulated peripheral loss was remarkably similar to the pattern found in a patient with real peripheral visual loss. Together, these findings point towards a possible perimetric screening test that is based exclusively on continuous gaze-tracking.

8.1.3 Eye-movement-based computational methods for visual field sensitivity mapping

Following up on the results from Chapter 3, Chapter 4 presented a study in which we developed two computational models to link oculomotor positional deviations to retinal sensitivity. These methods are (1) spatio-temporal integration of oculomotor positional deviations by means of Threshold Free Cluster Enhancement (TFCE) and (2) training a Deep Recurrent Neural Network (RNN) using the data acquired with the simulated scotomas from Chapter 3. These methods have complementary qualities: the TFCE is neurophysiologically plausible and its output correlates well with Standard Automated Perimetry, while the RNN significantly outperformed the TFCE in reconstructing the simulated scotoma with remarkable accuracy, but did not translate as well to the clinical data from glaucoma patients. Each of these methods requires further optimization, but both show there is potential for a faster, more intuitive alternative to Standard Automated Perimetry.

8.1.4 Motion sensitivity assessment based on the spatio-temporal properties of eye-movements

Chapter 5 presented a study on the relationship between motion sensitivity and the spatio-temporal properties of eye-movements. The oculomotor features that better correlate with changes in stimulus speed are the spatial and temporal uncertainties, i.e. the standard deviations of the Gaussian fit to the cross-correlogram between the velocities of the eye and the stimulus. The effect size of both features was $\eta_p^2 < 0.90$). This robust modulation of the spatio-temporal properties of eye movements did not correlate with the outcome of a conventional test of motion sensitivity based on Random Dot Kinematograms. This indicates that a gaze-based continuous psychophysics test to assess motion sensitivity can provide an observer-friendly way to assess aspects of motion processing that conventional trial-based psychophysics do not capture.

8.1.5 Oculomotor assessment of MS and PD patients based on a continuous gaze-tracking standardized test

Expanding upon the methods presented in Chapter 2, Chapter 6 described the development of SONDA: the Standardized Oculomotor and Neurological Disorders Assessment. This method provides a thorough characterization of ocular motility that includes the novel spatio-temporal properties and saccadic frequency distributions, as well as the standard parameters of the saccadic main sequence. The study cohort consisted of patients affected by Multiple Sclerosis and Parkinson's Disease. Each patient's measure was expressed with a z-score obtained by normalization with respect to a dataset of healthy, age-similar, controls. We showed that the features obtained with SONDA can be used efficiently with machine learning to classify the underlying neurological condition. Furthermore, we presented six cases of patients (selected out of 21) whose eye movement abnormalities were insufficiently captured by the conventional saccadic main sequence, but that the SONDA parameters did capture.

8.1.6 Attentional modulation of visual spatial integration: psychophysics and modeling

In Chapter 7, we modeled the relationship between spatial integration of visual information on the one hand and space-based and feature-based attention on the other. Using a hybrid psychophysical paradigm that combined continuous visual search interleaved with a two-alternative forced-choice paradigm, we found an interaction between spatial integration, quantified as peripheral crowding strength, and attention. Depending on the type of attention, crowding strength is modulated as follows: spatial attention induces a marked and selective reduction in crowding strength that is spatially consistent with the locus of attention, while feature-based attention induces a milder, but more widespread reduction of crowding strength throughout the visual field. Furthermore, expanding on previous work[12], we developed a population coding model that describes the mechanism underlying the attentional modulation. Starting from neurobiologically plausible parameters, the model predicts the strength of spatial integration of a neural population by adjusting the integration weights between nearby neural populations, with these weights being defined by the type of attention employed. The performance of the model is similar to that of human observers.

8.2 Discussion

In the introductory chapter of this thesis I posed two questions: 1) *How can spatio-temporal integration be modeled quantitatively?* and 2) *How can spatio-temporal integration properties be applied in a clinical context?*

In the following sections, I answer these questions in the light of the experimental results presented in Chapters 2, 3, 4, 5, 6 and 7, and I will outline possible future developments.

8.2.1 Beyond trial-based paradigms: on the theoretical and practical benefits of adopting continuous psychophysics to model spatio-temporal integration

A recurring theme throughout this thesis is the adoption of continuous experimental paradigms. For instance, Chapter 2 introduced a family of algorithms to extract spatio-temporal properties from continuous gaze tracking data, while Chapter 7 utilized continuous visual search with an attentional load to explore the mechanisms underlying spatial integration.

Traditionally, however, psychophysics and behavioral tests have been based on a trial-by-trial approach where the rigor of experimental variable control and apparent simplicity of analysis outweigh the need for ecological validity of the tasks used to investigate the brain.

The trial-by-trial approach relies on assumptions that are not always verified and validated by the researchers. The standard assumptions are the statistical independence of one trial from another and their identical distributions, although natural behavior is rarely discrete and the brain activity follows multiple time courses that do not necessarily obey a rigid trial structure[27].

To account for this potential discrepancy, a body of psychophysical studies in the last decade introduced the concept of serial dependence[212–214]. It can explain elements of neural activity that are otherwise labeled as "unexplained variance" or "noise", such as slow, spontaneous fluctuations of brain activity[215]. Although somewhat controversial[216,217], this effect has been found in many contexts, ranging from simple perceptual phenomena such as orientation[214,218], numerosity[219], position[220], and face perception[213], to more complex and abstract phenomena, such as judgments regarding statistical variance[221] and confidence[222].

As the world around us is substantially stable, the brain can utilize the temporal redundancy in the signals acquired from a visual scene to optimize its processing strategy, leading to desirable effects such as reduced cognitive load and response times[223]. This process is known as serial dependence. It is modeled as a weighted sum between stimuli adjacent in time, which is remarkably analogous to the weighted sum of stimuli adjacent in space present in other types of optimality models such as the feature integration model by van den Berg and colleagues[12] upon which I built the attentional integration model presented in Chapter 7. In both cases, the integration is modeled so that the resulting percept is substantially different from the sum of its spatial or temporal parts.

In the light of these considerations, moving beyond an exclusive trial-by-trial approach towards continuous paradigms seems reasonable and desirable for three reasons: 1) less time "wasted" in experimental downtime phases such as the inter-trial intervals; 2) paradigms that record a discrete behavioral response provide only an indirect explanation of the time-varying, noisy internal processes that the paradigms often intend to investigate; 3) perceptual phenomena that require the processing of more than the initial impulse of a stimulus can become more accurate with additional time, and would thus benefit from continuous assessment[27].

At this point, the (perhaps skeptical) reader might wonder about the applicability of what has been discussed so far to the context of eye movements, as they constitute a very prominent aspect of this thesis. The doubt is legitimate: some oculomotor components, such as saccades, notoriously show a stereotyped behavior (at least for individual saccades)[9], and how this relates to the concept of serial dependence across longer periods of time is far from clear[224–226]. The stereotyped nature of saccadic eye movements is at the core of the main-sequence analysis, an exponential relationship that ties the amplitude of a saccade to its peak velocity[121,122]. In Chapter 6, I showed how the main sequence, typically obtained with repeated trials of individual saccades, is not always the optimal characterization to detect eye movement abnormalities in Multiple Sclerosis and Parkinson's Disease. On the other hand, the spatio-temporal properties obtained with the continuous tracking paradigm introduced in Chapter 2 and extended in Chapter 6 were shown to be more sensitive to subtle alterations and were consistent enough to often enable the classification of the neurological disorder underlying the oculomotor impairment. A possible explanation for this is that tracking paradigms like SONDA (Standardized Oculomotor and Neuro-ophthalmic Disorders Assessment, see Chapter 6) are ideal candidates to account for serial dependencies[27], and thus to capture lesions and perceptual abnormalities that progress on completely different time scales than that of individual discrete trials.

There are multiple examples of tracking paradigms (both manual and ocular) that model the observer's behavioral response as a time series and that can account for serial dependence[2,26,227–229]. My paradigm was inspired by some of them, especially those used in the studies of Mulligan and colleagues[2,26] and Bonnen and colleagues[2,26]. It relies on a Kalman filter applied "in reverse" to estimate the unknown observer's noise variance. According to this framework, the observer uses knowledge about the target's dynamics (i.e. velocity and acceleration), the current (noisy) observation of target's position and a history of previous estimates of target positions to obtain an optimal estimate of the true target position at each time step. How these values are combined depends on the relative size of the two sources of variance present in the Kalman filter: (a) the observation noise variance and (b) the target displacement variance. The former is the variance associated with the current sensory observation (related to the observer's perceptual capabilities), the latter is the variance driving the target position from time point to time point (i.e, stimulus speed, visibility, anything that makes the stimulus easier or harder to track).

Normally, a Kalman filter combines these values (the noisy observation, target dynamics, previous estimates and sources of variance) to yield an optimal estimate of the current target's position, in order to approximate the actual, but unknown, position. In our case, however, the real target position is known. Therefore, the Kalman filter is "reversed" in order to estimate the observation noise variance, which serves as a proxy of the observer's perceptual capability.

In the context of oculomotor clinical assessment, this observer's noise can have multiple sources: perceptual (faulty input for eye movement command), neural (eye movement motor planning) or muscular (eye movement motor execution), which at present I was still unable to formally disentangle. However, the "reversed Kalman filter" estimate is only one of ten spatio-temporal features that I combined to quantify the oculomotor system of an individual. Others, such as the parameters of the Gaussian fit to the cross-correlogram between eye and stimulus velocity or the distribution of gaze-stimulus deviations, might be able to capture the complexity and peculiarity of each type of oculomotor impairment. In the future, the integrated approach of SONDA – unlike a trial-based approach alone – may be able to precisely identify the origin of various disorders. A case in favor of an integration of trial-based and continuous behavioral assessments is presented in Chapter 6, where the best oculomotor assessments were achieved by combining the discrete saccadic main-sequence with the novel, continuous, spatio-temporal properties. The implementation of the spatio-temporal properties in clinical assessment is addressed in the next section.

8.2.2 Beyond fixations and saccades: towards a clinical implementation of the spatio-temporal properties of eye movements

A large part of my work focused on the development of novel techniques to perform clinical evaluations relying on the spatio-temporal properties of eye movements (Chapters 3, 4, 5 and 6). I aimed to accomplish this while keeping in mind the perspectives of the people involved in clinical assessment – patients and clinicians – and without losing sight of scientific validity. I wanted to create a test that is fast yet comprehensive, easy to administer, understand and perform, and that is based on rigorous neurophysiological and mathematical foundations. A simple dot moving along a random path (or, "the drunken dot" as one patient once called it), turned out to be ideal for several practical reasons: parameters such as luminance contrast, speed and, size are easy to manipulate; it involves exclusively low-level perception, which makes it less susceptible to potential biases from culture or education; it involves continuous assessment over time (see the previous section); and, as stated before, it is very intuitive and fast (often the eye-tracker calibration took longer than the test itself). An assessment based on traditional parameters such as the fixation period or the number of saccades cannot fully utilize the potential of such a simple stimulus. In contrast, many studies have shown that a better and more flexible stimulus-response mapping is achieved when analyzing pursuit eye movements as behavioral time-series[230–235]. By using a framework that models sensory information processing as a noisy time-varying process[27], I not only created novel sensible and sensitive parameters (as in SONDA, Chapter 6), but also established a link between eye movements and retinal sensitivity. This unconventional approach to visual field assessment (perimetry) can potentially improve upon a number of critical aspects present in the assessment of ocular neuropathies like glaucoma. The current gold standard – Standard Automated Perimetry (SAP) – is often too taxing for such patients. As a result, SAP cannot be used with adequate frequency to correctly monitor their disease progression[236–239]. To detect visual field defects, other researchers have also attempted to create eye-movement-based perimetry[51,66], or have implemented a machine-learning approach[36]. However, the authors of the latter approach stated that it "offers no real hint about what particular characteristics of the resulting saccade maps are suggestive of abnormality". In contrast, my test relies on a physiologically plausible explanation of the relationship between visual loss and eye-movements: a non-linear combination of the magnitude and duration of gaze deviations from a target is proportional to retinal sensitivity. Essentially, large but transient errors and small but prolonged ones can both indicate visual field loss (see Chapter 4). Glaucoma and visual field defects, however, are not the only feasible applications of the spatio-temporal integration properties of eye-movements. As I showed in Chapter 6, neurological conditions that involve motor disorders are also suitable candidates. An especially interesting (and complex) one is Multiple Sclerosis (MS). It is an heterogeneous disorder in which the symptomatology depends on the location of the lesions in the CNS, with the most common oculomotor disorders being saccadic dysmetria[126]. This is present especially in patients with demyelination affecting the cerebellar peduncles[138] and pathological nystagmus on eccentric gaze caused by a failure of the common neural integrator of eye movements in the brainstem[114]. Quantitative eye movement assessments thus have the capacity to detect abnormalities otherwise missed during the clinical evaluation. This paves the way for the possible use of these tests to provide a behavioral biomarker that could detect disease progression and measure the effects of therapies. Early diagnosis of MS through eye movement would also be an important step, as timely initiation of immunotherapy seems to be the best strategy to arrest the progression of the disease[240]. Furthermore, it has been shown that eye-movement impairment in MS correlates with the scores of neuropsychological tests[241,242]. The use of eye movements to study cognitive processes is a well-established concept that has been around for decades[243,244], and has recently gained popularity as an additional tool for clinicians. A qualitative clinical examination performed by an expert neurologist can often be used diagnostically in disorders that primarily have motor symptoms but little cognitive impairment (such as Multiple System Atrophy[245]). Conversely, where cognitive impairment predominates over motor impairment, a quantitative assessment of eye movement characteristics is preferable (for instance in Alzheimer's Disease[246]). When both cognition and movement are impaired (for instance in Huntington's Disease[247] or Progressive Supranuclear Palsy[248]), a clinical examination is more useful during the diagnostic phase, while a quantitative assessment becomes essential to monitor disease progression over time[249,250]. The fast rise in advanced machine learning approaches such as Deep Learning, combined with high-throughput eye-movement tests such as SONDA (Chapter 6), Tseng and colleagues' test[146] or DEMoNS[123], will hopefully enable clinicians to integrate extremely efficient, non-invasive, eye-tracking tests into their daily practice, alongside their conventional confrontational examinations.

8.2.3 Concluding remarks

The overall conclusion that I draw from my work is that the spatio-temporal integration of visual information—when operationalized through eye-movement responses, motion sensitivity, and peripheral crowding magnitude—is a flexible approach that can potentially be used to diagnose ophthalmic (Chapters 3 and 4) and neurological (Chapter 6) disorders as well as to study perception (Chapter 5), and cognition (Chapter 7). Finally, investigating and modeling integration based on spatio-temporal properties has enabled me to understand the importance of time in Vision: knowing *when* we see something is as important as knowing *what*, and *how* we see the world around us.