Biologically encoded augmented reality: multiplexing perceptual bandwidths

by

Matthew Everett Lawson

B.F.A., School of the Art Institute of Chicago (2010)
M.S., Massachusetts Institute of Technology (2012)

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Media Arts and Sciences

at the
Massachusetts Institute of Technology
September 1st, 2020

Author………………………………………………………………………………………….
Program in Media Arts and Sciences
September 1st, 2020

Certified by ………………………………………………………………………………………..
Tod Machover
Muriel R. Cooper Professor of Music and Media
Thesis Supervisor

Accepted by ………………………………………………………………………………………
Tod Machover
Muriel R. Cooper Professor of Music and Media
Academic Head, Program in Media Arts and Sciences

Biologically encoded augmented reality: multiplexing perceptual bandwidths

by

Matthew Everett Lawson

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, on September 1st, 2020 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Media Arts and Sciences

Information floods the center of our visual field and often saturates the focus of our attention, yet there are parallel channels constantly and unconsciously processing our environment. Using sensory plasticity as a medium to create new visual experiences allows us to challenge how we perceive extant realities. This dissertation describes a novel system and methodology for generating a new visual language targeting the functional biology of far peripheral mechanisms. Instead of relegating peripheral vision to the role of redirecting attentional awareness, the systems described here leverage unique faculties of far peripheral visual processing to deliver complex, semantic information beyond 50° eccentricity without redirecting gaze.

Synthetic position shifts can be elicited when frequency gratings or random dot patterns are translated behind a static aperture and viewed peripherally, a phenomenon called motion-induced position shift (MIPS). By transforming complex symbols into a series of strokes articulated through MIPS apertures, I present a codex of motion-modulated far-peripheral stimuli. Methodologies describe a two-stage implementation: first, proven psychophysical constructs are integrated into contextual forms, or codex blocks, and second, adapted to an environment of complex scenes. This approach expands upon prior work not only in first principles of visual information composition and delivery but also in its capacity to convey highly abstract forms to the far periphery with no gaze diversion, via apertures spanning only 0.64 degrees of the visual field.

Spatial compression and far peripheral delivery of complex information have immediate applications in constrained display environments for interaction, navigation, and new models of visual learning. As the technological cutting edge outpaces our physiological sensitivities, the proposed methodologies could facilitate a new approach to media generation utilizing peripheral vision as a compression algorithm, redirecting computation from external hardware to direct correlations within our biology.

Systematic and applied longitudinal studies were conducted to evaluate the codex in increasingly complex dynamic visual environments. Despite increasing scene complexity, high detection accuracy rates were achieved quickly across observers and maintained throughout the varied environments. Trends in symbol detection speed over successive trials demonstrate early learning adoption of a new visual language, supporting the framework and methods for delivering semantic information to far peripheral regions of the human retina as valuable extensions of contemporary methodologies.

Thesis advisor: Tod Machover
Muriel R. Cooper Professor of Music and Media

2

Biologically encoded augmented reality: multiplexing perceptual bandwidths

by

Matthew Everett Lawson

This dissertation/thesis has been reviewed and approved by the following committee members

Tod Machover ……………………………………………………………………………………..
Muriel R. Cooper Professor of Music and Media
Massachusetts Institute of Technology

Paulo Schor ………………………………………………………………………………………
Director of Innovation and Professor of Ophthalmology and Visual Sciences
Universidade Federal de São Paulo

Joseph A. Paradiso ………………………………………………………………………………
Alexander W Dreyfoos (1954) Professor in Media Arts and Sciences
Massachusetts Institute of Technology

Kent Larson ……………………………………………………………………………………....
Director of City Science and Principal Research Scientist
Massachusetts Institute of Technology

Contents

For Linda Peterson,

one of the greatest minds and leaders of our time

Dedications
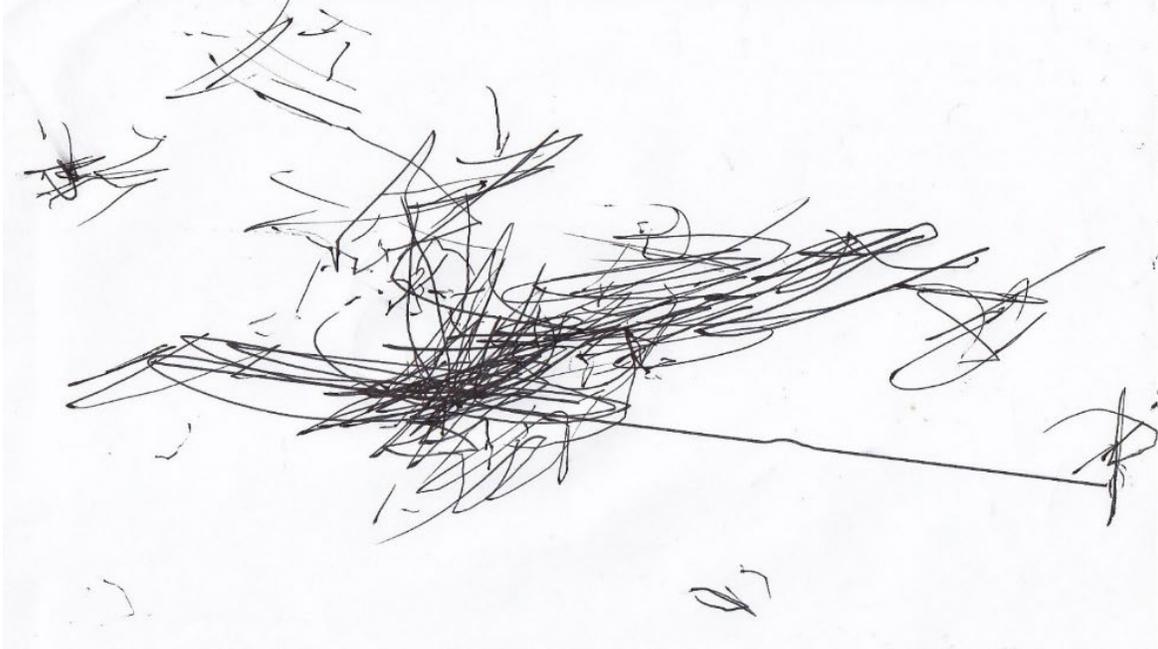
Figure 1. Venn diagram: visual perception (source: August Hawthorne Lawson, age 1)

I would like to acknowledge, with all of my heart and being, the innumerable people who have encouraged the world we have to build: my wife, Amy, my son, August, and the dedicated professionalism and shared humanity of my committee. Tod Machover is one of the most brilliant and generous humans I have ever had the pleasure and honor to work with. There are few people you get to meet in life that radically change the way you see. Dr. Paulo Schor has not only been my greatest mentor but has a perspective on the universe that I will spend the rest of my life just attempting to catch up to. The formalization of this research, which has been at the core of my curiosity could not have been better guided than by Joe Paradiso and his invaluable perspective on the scientific process. We all work between fields and I cannot be more grateful to Kent Larson for challenging me to see beyond my writings and notebooks, to imagine what the future can look like.

Without all of you, the tune I've been whistling my entire life wouldn't be true.

When you wish upon a star

Time to build a telescope.

1. Introduction

Figure 2. Boat

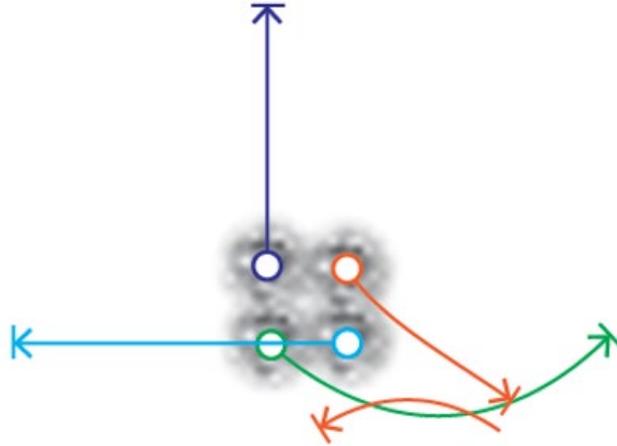Communication technologies and their potentials for extended application lend rich opportunities to activate receptive fields in the human visual system. Peripheral vision has long been targeted as an opportunistic space for its potential to expand available real estate for information delivery. Flicker, for example, is highly detectable peripherally, but as a binary system is ill-suited for alphanumeric characters or complex symbol cognition. Peripheral shape discrimination is feasible with similar sensitivities as in central vision but requires an increase in <u>shape size with increasing eccentricity from central fixation for accurate detection</u> [1]. This dissertation demonstrates a foundational approach to peripheral information delivery capable of conveying highly complex symbols via motion-modulated stimuli within <u>a series of small, static apertures in the far periphery (> 50°)</u>. Methodologies described here utilize perceived position shifts as a medium toward a new visual language, conveying complex information and representational characters well outside established limits of symbol recognition [2] while occupying comparatively little spatial real estate in both pixel space and retinotopic projections onto discrete regions of the retina.

Parsing specific channels of visual attention and understanding its bandwidth, boundaries, and limitations traditionally necessitate the use of carefully controlled environments [3].

These constraints have created a longstanding gap between potentially transformative discoveries in vision science and their application in naturally and dynamically changing real-world environments, which we navigate daily. This dissertation argues potentials for new media formats enabling *biologically* encoded augmented reality, through novel methods leveraging phenomena unique to early visual processing mechanisms, expanding dimensions of semantic information delivery.

I present a methodology for the development of a codex of experimental symbols, their computational throughput from genesis to observer, with twelve unique static-aperture motion-modulated envelopes, further referred to as *codex blocks* throughout this discourse. These blocks, when integrated into an observer's visual field at <u>far eccentricities</u> are shown to convey not only alphanumeric characters but also complex geometries and abstract symbols with high adoption rates for accuracy and recognition. These codex blocks were presented over longitudinal human studies to quantitatively assess rates of detection and accuracy in observers, and demonstrate <u>trajectories of visual learning in varied environments</u>. The findings represent a foundational approach to the programmatic engineering and implementation of novel systems to integrate peripheral motion semantics into dynamic real-world scenarios.

1.1 System overview

This research follows theoretical practices in basic histology and practical function of visual mechanisms to support the computational development of novel platforms. Chapter organization is designed to give readers a high-level introduction to relevant fields and their parts encapsulating contributions in this research. Each chapter can be treated separately; however, it is my intention to build a narrative for readers such they may follow a developmental throughput from first principles to applied study outcomes. There are many conceptual issues following practical application addressing conscious and unconscious visual processes, and albeit outside of the scope of this dissertation, the discussion appending will couch various perceptual micro genesis from representational to phenomenal, in addition to neurophysiological architecture commonly used to express a larger sphere in contemporary practice.

Unique characteristics of early visual processing yield a powerful effect called motion-induced position shift (MIPS) which enables the approach outlined in this dissertation. When frequency gratings or random dot patterns are translated behind a static aperture and viewed peripherally, the aperture itself appears to move. This effect has been parameterized in literature and hinges on a condition called "equiluminous kinetic edges," when the luminance span of patterns behind apertures are bisected by the luminance of the surround.

Leveraging abilities to create synthetic motion in the periphery, a new visual language is described transforming characters and abstract symbols into a series of strokes, functioning as motion paths for generating MIPS stimuli. Multiple apertures can be arranged to articulate multi-stroke symbols, laying the groundwork for implementing almost any conceivable character or symbol[1]. Studies were designed to evaluate first principles of this approach in a standard psychophysical setting in addition to assessing feasible implementation of these techniques in increasingly complex visual environments.

A simulated environment was constructed to administer studies, consisting of a display array surrounding an observer. In all study series codex blocks were presented at 55° eccentricity on either left or right display. Codex blocks in *control* series were overlaid on a 50-gray background with central fixation and in following *low load* and *high-load* series were overlaid on dynamic natural scenes. These study methodologies emulated visual experiences of driving or walking where <u>significant scene variations and foveal distractors occur over time</u>. Source footage was captured from two distinct points of view: a pedestrian while viewing a hand-held device, a pedestrian facing forward, and an operator of a vehicle. Participants were introduced to the codex over randomized series of presentations. They were instructed to indicate via button press as soon as they had high confidence in symbol recognition and their answer was recorded. Gaze tracking was captured to exclude any data outliers resulting from significant diversion in fixation.

---

[1] This system would of course be subject to the same challenges of distinguishing similar characters (i.e. 1 and l) as other forms of visual communication.

Activating far peripheral zones of visual fields for information delivery is motivated by several immediately relevant applications. Navigation and situational awareness tasks in cockpits under human control and assist are key examples of demanding environments with significant burden on visual attention, fixation, and cognitive load. Communicating sophisticated visual information would normally require some interruption of central visual field for reliable detection accuracy. Diverting more information to peripheral channels increasingly frees the central visual field for immediate tasks. Near eye displays, such as peripheral head-mounted displays and other augmented reality technologies are especially constrained in pixel space available for information delivery and would benefit greatly from future systems integration capable of condensing many bytes of information into small pixel form factors. Perhaps the most exciting potential of exploring the development of a new visual language is greatly expanding capacities for <u>comprehension and reading via far periphery</u>.

From how we navigate visually unpredictable environments with a multifaceted restriction on visual attention, to perceiving the made environment through architecturally augmented skins, to arguing for a deeper look into the way we generate media and display technologies, this dissertation will explore real-world application for unattended visual potentials. The outcomes of the proposed experimentation could have direct application in generating new models for learning, interaction, navigation, and begin to define emergent potentials in immersive environments.

2. Foundations

2.1 Biological vision

The architecture of a system to deliver carrier signals[2] coded into codex blocks onto an observer's awareness must be informed by fundamental biological systems that organize visual information spatially from the optics of the human eye to the distribution and density of rods and cones on the retina. We can explore spatial vision as if in a single frame of time-space, but most interesting, and impactful, physiological processes are those resulting in visual perception of motion over time. The following sections will move somewhat linearly through the human visual system, from the physiology of light capture on the retina, to organization of spatial information, to receptive field function and allocation, and to detection of image motion as it directly relates to this work.

2.1.1 Retinal image formation and visual fields

Physical organization of photoreceptive cells in the retina results in a segmentation of regions of the visual field: foveal (or central), parafoveal (or near-peripheral), and peripheral. Photoreceptors responsible for color detection are densely packed in the central 0.032 deg^2 of the retina, while monochromatic photoreceptors are completely absent from the central 1.25° before increasing superiorly and, less rapidly, nasally [4]. Other important visual capabilities vary across the visual field, though the cause of this functional segmentation is found further downstream in later more sophisticated visual processing mechanisms.

---

[2] The term "carrier signal" in the context of this dissertation refers to psychophysical structures, such as Gabor patches, which are adapted to environmental parameters and animated along symbol stroke paths

Table 1 summarizes some functional segmentations given by their deviation from line of sight (LOS).

Table 1. Visual function varying across segments of the human visual field [2].

Visual function	Angular deviation from LOS (degrees)
Word recognition	±1°
Symbol recognition	±30°
Color discrimination	±30° to ±60°

Further studies examining limits of peripheral symbol recognition have found that accurate identification of *characters* is contrast dependent and at 100% contrast maximum fields of recognition can be extrapolated to 46° eccentricity from central fixation [5].

One prominent hypothesis concerning differences in structure and function between foveal and peripheral vision is that one is a scaled version of the other – spatially, temporally, and photometrically. This is supported by findings that differing distribution of ganglion cells (more densely packed in the fovea and sparser in the periphery) result in grating sensitivity and Vernier acuity measurements in fovea and periphery differing by a scaled factor [6], [5]. This is also referred to as cortical magnification, or M scaling. Figure 3 depicts regions of the retina and their corresponding cortical scale.

M scaling alone, however, does not account for differing capacity for complex stimuli recognition between foveal and peripheral vision [7], [8]. Shapiro et al have hypothesized peripheral visual systems also combine features foveal visual system can separate, resulting in a "feature blur" and other spatiotemporal discontinuities between the two visual fields [9]. While scaling and blurring of features renders peripheral vision less suited to parsing static shapes and symbols, these features are the foundation of motion-induced position shift effects, which are the basis for stimulus delivery mechanisms at the foundation of this dissertation.
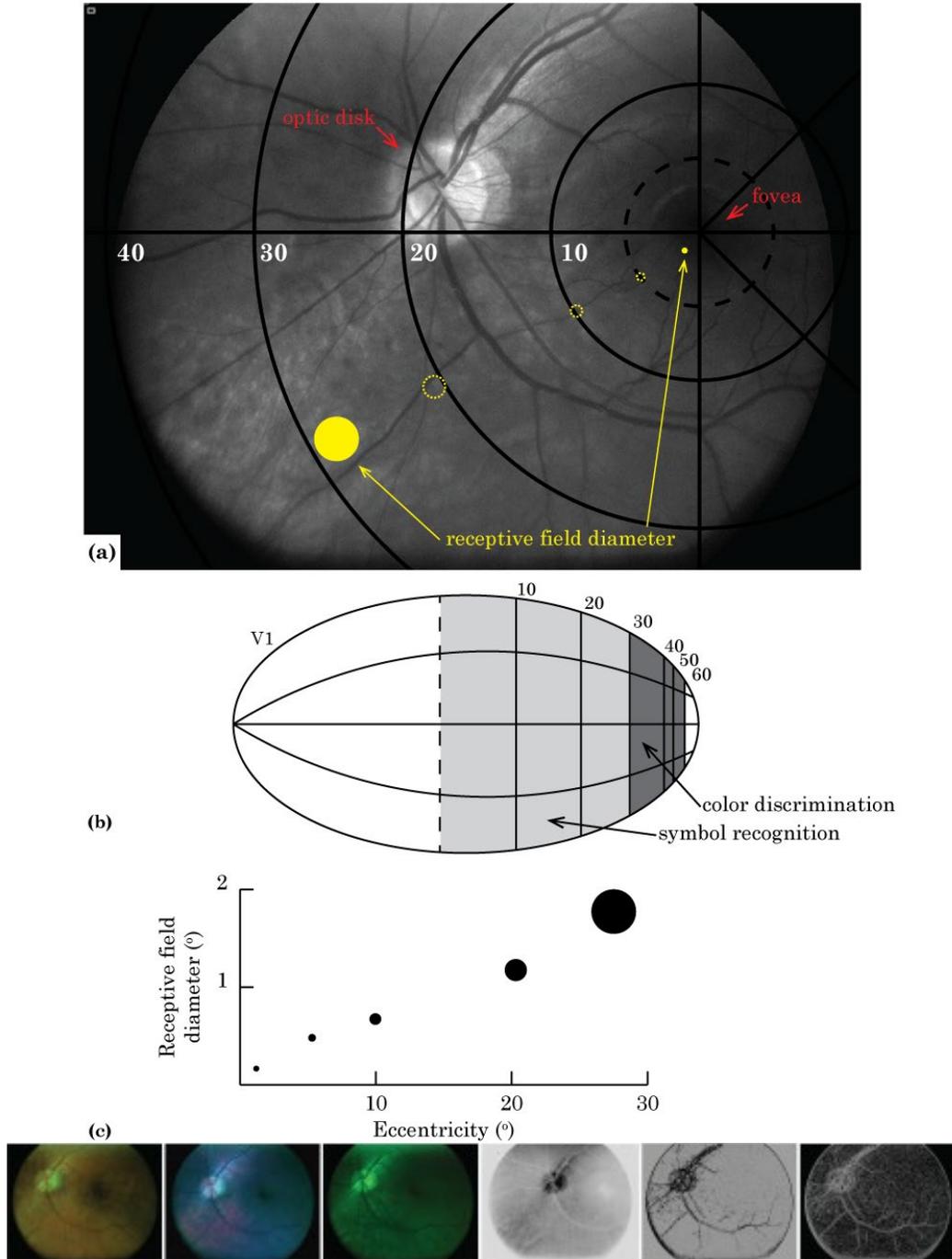
Figure 3. Cortical magnification given by scaled receptive field size at increasing eccentricities. (a) A retinal image of my left eye with annotations indicating notable structures in red and relative receptive field size with increasing eccentricity in yellow. (b) A representation of scaled cortical size for each delineated eccentricity. (c) Receptive field diameter as a function of eccentricity. Further images below of my left retina captured under different lighting conditions. Prototype imaging device specifications can be found in the appendices.

2.1.2   Receptive fields as early edge detectors

Information encoding and translation through receptive fields and ultimately toward a state of perception is both fascinating and daunting in its nearly inextricable complexities but is important to bring into form as these very mechanisms of spatial frequency, orientation, amplitude, contrast, and phase play a fundamental roll in the design of experimentation throughout this dissertation.

Field and function of ganglion cells, for simplicity's sake, can be thought of as electromechanical binary chemical gates:
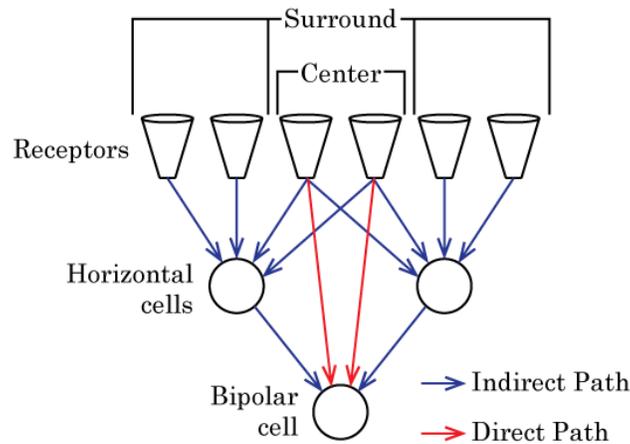
Figure 4. Lateral inhibition between early receptors in vision. Photoreceptors send a graded output signal to bipolar and intermediate horizontal cells which in turn complete the indirect path to the bipolar cell. This results in antagonism between the two pathways. Adapted from [10].

For example, when the center of *on-center* cells are stimulated they produce a spike discharge but when the center of *off-center* cells is stimulated, response is absent until stimulus is removed. This antagonism between the inner circle and surrounding annulus is called lateral inhibition – capacity of an excited neuron to disable spreading of action potentials from excited neuronal fields to their neighbors. Functionally, this means that receptive fields reduce cross-talk and noise at a cellular level and act as localized filters that streamline compressed data further along sensory perception pipelines.

These initial receptive field signals continue to cascade through the visual system to the lateral geniculate nucleus (LGN), which is like a weigh station for low-level processing, responsible for cataloging and preserving spatial information from the retina. This is referred to as "retinotopic mapping," a crucial function allowing the LGN to continue signal filtering and masking reducing raw external stimuli to fundamental features, both in the

spatial and temporal domain. This same retinotopic mapping is the basis for M-scaling mentioned earlier and contributed a crucial role in informing the design of stimuli within the codex apertures, exciting these pathways and yielding new perceptual feedback mechanisms.

In addition to eccentricity dependent scaling, receptive field sizes also change along the ventral stream, an area of the visual cortex responsible for discerning spatial patterns, objects, and scenes [11]. Population receptive fields (pRF) represent groups of neurons that respond to a visual stimulus [12] and fMRI estimates of pRF eccentricity scaling throughout the visual cortex agree with electrophysiological measurements[13].

This pipeline from photoreceptors to V1, ultimately culminating in what I would call a biological potential for stabilizing our relationship with the external visual environment, is the foundation of the mechanisms underlying many crucial filtering stages in early visual processing, including perceptual motion.

2.1.3   Conceptualizing motion

We could imagine all perceptual and psychophysical experimentation to have been performed with a static observer, eyes and

head fixed in space and gaze, but there is a nontrivial challenge to understand motion in dynamic real-world scenarios. The eye can sweep across a static scene and produce image motion on the retina, likewise track a moving object with the intraocular muscles, head, body in almost an innumerable complication of geometrical vectors while maintaining a relative perceived understanding of the underlying physical environment. The mechanisms of how a dynamic optical event on the retina manifests in a perception of a moving object are still not fully described by psychophysical or neurobiological explorations. Expressed simply, the visual system achieves this in two parts: the early analysis through the registration of 2-D image motion and further processes of interpreting image motion in terms of 3-D objects moving in 3-D space.

The starting point for motion and perception is what Gibson called the dynamic ambient optical array: the flux of optical structures converging from all directions to a station point [14]. Foundationally it has been mapped and is known as 5-D plenoptic function [15].
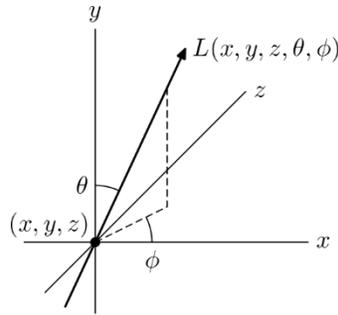
Figure 5. The 5-D plenoptic function [16]

What is undisputed, however, is there are directionally selective cells in the visual cortex that play a crucial role in motion perception. In most computational theories image motion is based on some version of a delay and compare method [17], where "what happens" in one region of the retina integrates with "what happened" shortly before in nearby regions, with the assumption of similarities between neighbors (similar signals). There are also edge-based models [18] suggesting that analyzing the change in illumination over time in conjunction with an edge detector could explain motion.

These models sufficiently contextualize simple motion constructs. Some visual phenomena, however, create a discordance between actual and perceived motion. These perceptual motion effects represent some of the most potentially impactful stimuli within the context of this dissertation. Recent studies by Schneider et al. show such illusory motion shifts result in actual measurable activation of neighboring receptive fields [19], similar to the activation spurred by actual motion stimulus.

Ultimately the class of motion models most relevant to this work is spatial frequency-based techniques, which use filters constructed from simple building blocks analogous to isolated mechanisms within the visual system. These filters can be combined with space-time projections of motion stimuli to extract spatiotemporal energy, a very useful analysis for contextualizing motion, whether actual or illusory [15]. This motion energy analysis is used to quantitatively evaluate the twelve unique codex blocks used throughout the observer trials.

2.2 Psychophysical constructs

There are several fundamental constructs, refined over decades of research elucidating the causal links between physical stimuli and the manifestation of visual perception, comprising the building blocks of the approach outlined in this dissertation.

Figure 6. Standard Gabor patch.

Gabor patches have been used extensively within vision science research, favored for their simple analytical expression and ease of modification [20], and studies demonstrate Gabor patches mimic 2D spatial responses of simple cell receptive fields [21]. It is essentially a one-dimensional sinusoidal luminance grating that is Gaussian windowed in space, and is given by the following general expression:

$$L(x,y,t) = L_m\left\{1 + C_p cos[2\pi x f_c + \theta_c(t)] \times exp\left[-\frac{1}{2}\left(\frac{x}{\sigma_x}\right)^2 - \frac{1}{2}\left(\frac{y}{\sigma_y}\right)^2\right]\right\}$$

Where L_m is mean luminance of the display, C_p is peak contrast of the Gabor, f_c is the grating spatial frequency (the carrier), and σ_x and σ_y are standard deviations of the spatial Gaussian window. The cosine term gives sinusoidal grating and the following exponential term generates the Gaussian window.

White Gaussian noise (WGN), when windowed and blurred with varying standard deviations, is also an impactful tool in targeting visual processes responsible for motion detection. A 1D Tukey filter [22] as defined below is used throughout studies when preparing stimuli to create a window with sinusoidal falloff:

$$w(x) = \begin{cases} \frac{1}{2}\left\{1 + cos\left(\frac{2\pi}{r}\left[x - \frac{r}{2}\right]\right)\right\}, & 0 \leq x \leq \frac{r}{2} \\ 1, & \frac{r}{2} \leq x \leq 1 - \frac{r}{2} \\ \frac{1}{2}\left\{1 + cos\left(\frac{2\pi}{r}\left[x - 1 + \frac{r}{2}\right]\right)\right\}, & 1 - \frac{r}{2} \leq x \leq 1 \end{cases}$$

2.2.1   Multidimensional Gabor arrays

Further complexity in Gabor-structured stimuli has been demonstrated by tiling multiple patches into an array presented foveally. When multiple Gabor patches are oriented in an array the resultant effect is determined by the relationship between orientation and motion of individual patches. Carrier signals with randomized θ (orientation) and differing velocities result in independent position shifts, normal to the angle of orientation of the carrier. However, if drift velocity is consistent across the array, a pooling occurs resulting in the sensation of global motion of a coherent surface [23], [24].

Thurman and Lu used a such an array to evaluate the ability to distinguish biological forms, decomposing motion of a walking figure into discrete Gabor discs, oriented and oscillating according to gait mechanics, as if a stick-figure made of sinusoids was walking in place behind a wall of circular windows [25]. They found observers were capable of distinguishing biological motion from incongruent cues, an exciting implication for the capacity of Gabor arrays to convey sparsely sampled dynamic objects. While exciting, these findings do not apply directly to the class of visual stimuli used in the approach described in this dissertation because they are (1) viewed foveally and (2) perceived as an array of apertures with a continuous signal moving behind, which is fundamentally <u>contrary to motion-induced position shift mechanisms described below.</u>[3]

2.3 Motion-induced position shift (MIPS)

When Gabor patches are shifted behind an envelope over time and positioned in the far peripheral region of the visual field, it is possible to produce a perceived position shift though the envelope remains fixed spatially. These effects are significant because such motion illusions defy the concept of a precise retinotopic localization of the striate cortex – in other words, because mapping of retinal to cortical regions is so exact, if the stimulus is not actually

---

[3] There are two other fascinating examples of global motion from multidimensional Gabor arrays that, while not directly related to this work, are worth mentioning here: "Motion drag induced by global motion Gabor arrays" [83] and "A comparison of global motion perception using a multiple-aperture stimulus" [84].

moving and activating neighboring regions on the retina, then the signal causing perceived motion must reflect a "re-entrant mechanism" of feedback from V5 to V1 [26][27][28]. By adapting the underlying first principles of the Gabor effect and expanding beyond a simple aperture and grating, I propose generating new functional perimeters. This dissertation presents a system methodology, leveraging motion-induced position shifts as a tool for delivering highly semantic information peripherally. These geometries and mechanics act as an inverse scribe of a natural scene presented to the visual pipeline, activating new modes of media delivery in dynamic environments.

2.3.1   Necessary conditions for successful MIPS and shift magnitude quantification
Successful activation of peripheral mechanisms yielding an illusory position shift is contingent on a set of stimulus characteristics. De Valois and De Valois [26] illustrated that "movement-related bias" was most pronounced when:

a.   at greater eccentricities within the visual field (at a rate of 1-2 arcmin per degree of eccentricity),

b.   at 4-8 Hz temporal frequency and at low special frequencies, and

c.   when motion of the pattern behind the aperture was moving toward or away from the fovea rather than tangentially.

More recently, it was found that most pronounced illusory shifts seem to occur regardless of spatial frequency of the carrier signal (moving stimulus behind an envelope) with maximum position shift achievable at both intermediate [26] and low [29] spatial frequency. Rather, the relationship of luminance of the carrier signal to the surround is especially critical to the success of the effect: equiluminous kinetic edges, occurring when luminance values of the carrier match the background, cause the observer to perceive the patch as a "voluntary figure" moving independently from the surround. With non-equiluminous kinetic edges, the envelope appears as a hole through which the carrier is observed to be moving and the effect collapses [30]. Mechanisms that produce such "kinetic edges" are entwined within the complex manifestation of vision, but recent studies have shown that MIPS is even perceptible dichoptically, suggesting that early monocular processing contributes significantly to the effect [31].

One of the most interesting implications of the motion-induced position shift effect is *perceived area* covered by stimuli and how it can greatly expand perceptual real estate within the visual field with no change in envelope size. This perceptual area expansion is non-uniform and varies with the velocity of the carrier signal. At low velocities, the magnitude of the effect increases monotonically with duration, but as velocity increases magnitude decreases before reaching a steady-state [32]. The relationship between the velocity of carrier frequency and perceived position shift also varies with the type of carrier signal: first- or second-order. Perceived shift for a first-order, luminance based Gabor patch increases with carrier velocity, but only to a point – the velocity cap appears to occur at approximately 1 deg/s [33]. Bressler and Whitney found that, for second-order carrier signals, temporal frequency dependence is more of a band-pass than the high-pass characteristic of first-order stimuli [29], [32]. This inherent coupling of velocity and effect magnitude, if non-linear and somewhat complex, is a critical determinant in symbol path design.

2.3.2   Increasing MIPS signal complexity for dynamic environments

At least two[4] dominant theories of motion perception have emerged, distinguishing visual mechanisms by the class of stimuli to which they are sensitive. First-order motion detection is responsive to changes in *luminance*, and computational models of this architecture were proposed first by Hassenstein and Reichardt (Reichardt detectors)[34] and then by Adelson and Bergen [15]. Second-order motion was subsequently classified to address a subset of visual stimuli not detected by first-order models [17]. While first-order motion is manifest by changes in luminance, second-order motion is responsive to stimuli that are equiluminant, or with even luminance across field and changes only in contrast or spatial frequency.

Motion-induced position shifts were first illustrated with rather simplistic first-order motion stimuli, but further works show effects are extendable to second-order motion as well. The first demonstration utilized a square aperture and random dot pattern (first-order, or

---

[4] The systems of functional organization described here are included because they appropriately contextualize the types of stimuli deployed in this dissertation, but by no means provide a complete classification.

apparent motion) [30] and was later modified to use a dynamic Gabor (second-order, or constant Fourier energy motion) [29]. The extension of proven equiluminous kinetic edges with varying motion complexity greatly expands the applicability of perceived motion shifts to increasingly dynamic environments.

2.4 Visual attention and perceptual load

While the formation of an image on the retina is the foundation of visual processing, it is not entirely determinant of perceptual outcome. The field of visual attention research has grown significantly, gaining in prominence as it became evident that changing an observer's "attentional state," while the image formed on the retina remains constant, can affect activity of sensory neurons throughout the visual cortex [35].

Implementation of systems described in this dissertation necessitates an understanding of how the human visual system allocates attention and awareness within the visual field. Further, beyond how this is accomplished within short periods (in the range of hundreds of milliseconds) it is important to understand how attention becomes saturated fluctuating over time. Assuming the amount of attention available at any given time is finite lends toward a theory of perceptual load [36] which allows us to incorporate a quantitative assessment of capacity within the metaphor of attentional selection.

The first important implication of this theory is perceptual load present at any given time is a determining factor for how and when selective attention occurs. More difficult tasks (with high load) play a role in selective attention by prohibiting interference from irrelevant information [37]. This interference can be so complete with sufficiently high load tasks that this effect has come to be known as "inattentional blindness" [38]. Further studies have clarified the extent of this effect by using functional imaging during tests of both meaningless and meaningful distractors presented during high load tasks [39]. Results of such experiments illustrate that brain activity does not differentiate between two distractors during high load tasks (even when looked at directly), implying that visual recognition necessitates attention even for highly meaningful stimuli.

Recent studies, however, demonstrate that increases in attentional load do not affect perception of double-drift stimulus, a related position shift effect [40]. From these studies we can expect that, while peripherally presented semantic information may not overlap spatially

with central visual field, an attentional occlusion may occur during symbol recognition. The goal of studies implemented throughout this dissertation was to quantitatively evaluate codex block detection speed and accuracy in environments of varying complexity, as a foundation for deploying this approach in dynamic scenarios. Important next steps in future research will fully assess attentional cost of any multiplexed information delivery via true cross-task procedures measuring observer performance on an attentional task while perceiving codex blocks. (see 7.5 Limitations)

3. Contributions

This dissertation demonstrates a foundational approach to peripheral semantic information delivery capable of conveying highly complex symbols well beyond the established mean, using motion-modulated stimuli within <u>a series of small, static apertures in far periphery (> 50°)</u>. It describes novel system architecture inspired by an opportunity to integrate powerful programmable phenomena onto an observer's perception in complex visual environments. By leveraging proven psychophysical constructs, such as motion-induced position shift, systems described in this chapter translate characters and abstract symbols into motion stimuli, or *codex blocks*, and deliver them to far-peripheral zones, challenging limits of complex symbol recognition in the literature.

The system architectures described here represents a new intersection of the fields of vision science, computational imaging, and display technologies and could challenge the way we generate media for human consumption in active environments. In summary, this research expands upon prior work by presenting the following contributions:

1. Translation of alphanumeric and abstract symbols into motion-modulated static aperture stimuli which:
   o are detectible with high accuracy rates at eccentricities greater than 50° from fixation
   o occupy small regions of the visual field (less than 1 degree)
   o are scalable to symbols containing multiple strokes for communicating increasingly complex semantic information
2. Adaptation of visual stimuli for deployment in complex natural environments, such as from a pedestrian or automotive cockpit point of view.

The unique advantage of this approach is compression of highly semantic visual information into small static apertures, overcoming cortical scaling to efficiently communicate characters and symbols to receptive fields at <u>greater than 50° eccentricity</u>. Furthermore, a set of translational methods is described for deploying peripheral stimuli outside of controlled laboratory conditions.

Perceptual shifts achievable with equiluminous kinetic edges, multidimensional or otherwise, are traditionally established, evaluated, and utilized in standardized psychophysical environments (i.e. in a darkened room, viewing screens with controlled luminance ratios between stimulus and surround)[3]. Studies evaluated the persistence of MIPS in dynamic natural environments absent of static fixational cues while preserving as much as is possible the requisite relationships between stimuli and surroundings. Longitudinal studies conducted across observers evaluated the feasibility of deploying codex blocks in increasingly complex visual environments by recording symbol detection times from within a <u>standardized control series in a psychophysical environment, to pedestrian and automotive cockpit point of view (POV).</u>

This chapter will first address: how codex blocks target and activate discrete visual processes. Considerations in the design of this new structural form are summarized, including methods of altering symbol stroke path geometries to maximize perceived position shift and positioning multiple MIPS apertures for multi-stroke symbols. Second: processes of adapting codex blocks for integration within dynamic natural scenes are detailed, followed by descriptions of systems implementation for experimentation.

3.1 Translating motion to meaning

For the sake of simplification, we can think of any object static or unchanging in our periphery as invisible through habituation. Motion is the catalyst to perception in the discrete outer regions of the retina. A simple way to overcome imperceptibility of static forms in the far periphery is to apply <u>a pattern consisting of Gaussian local motion</u> to a character or symbol, such as Schäffel's "Luminance Looming" effect [41]. However, presenting symbols at increasing eccentricities in the visual field must account for cortical scaling, which at far peripheral locations beyond 40° would exceed a factor of 3 [5]. For the application of peripheral information delivery, this factor is highly impractical. To overcome cortical scaling, semantic forms are parsed into strokes and conveyed through motion-induced position shifts from small, static apertures. In the viewing environment developed for experimentation, aperture diameters were set at 35px, or 0.64 degrees of observer's visual field. Utilizing multiple MIPS apertures for multi-stroke symbols, this approach can theoretically articulate almost any conceivable character or symbol in fractions of a degree of the visual field.

As a demonstration of first principles, twelve unique symbols were generated and transformed into motion-modulated static aperture codex blocks, characters: A, B, C, D, 2, 3, 4, 5, in addition to abstract *forms* such as boat, stick figure, flower, and tree. Each symbol was first parsed into its constituent stroke paths (see Figure 7 example of character *3*). An image of white Gaussian noise (WGN) was animated along the stroke path behind a Gaussian windowed aperture to generate a perceived position shift. Grain size of WGN patches was selected based on aperture diameter, such that sufficiently unique features in grain were visible relative to their Gaussian fall-off as the patch translates behind the aperture [42].

Symbol parsing and stroke design described in Figure 7 yields a perceived position shift with a magnitude relative to the cumulative area in pixel space occupied by the 35px diameter envelope when animated along a motion trajectory. Chart (C) demonstrates several orders of magnitude increase in resultant perceptual size when compared to actual pixel real estate. These calculations are approximate and do not account for variations in perceived stroke magnitude due to varying velocities or trajectories. An immediate avenue of future work is to quantify these variations across observers.
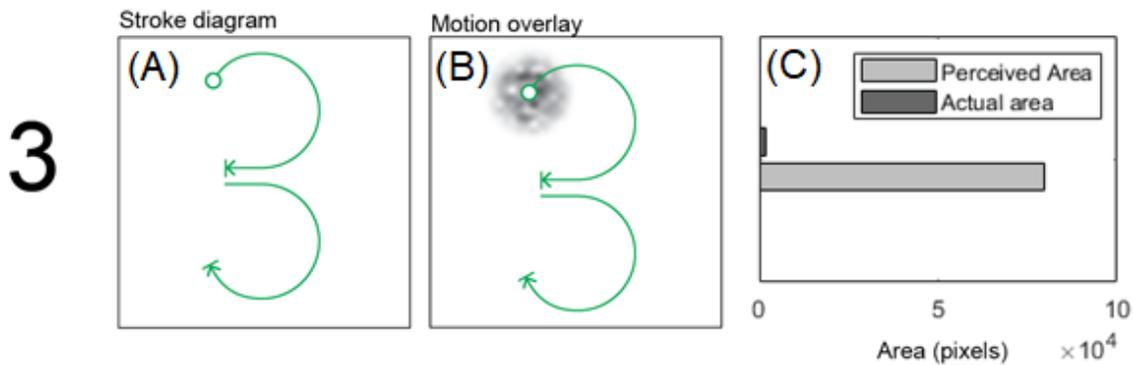
Figure 7. Example of parsing a character into strokes conveyed through motion-induced position shifts behind a static aperture. (B) Stroke paths in green show the path followed by the WGN pattern as it moves behind the aperture. (C) Perceived versus actual area is shown. Actual area is given by the diameter of the static aperture (in this example, 35 px). Perceived area is derived from the cumulative area covered by the 35px aperture as it traverses the motion path.

Stroke paths were engineered for optimal perceptibility given known parameters of motion-induced position shift phenomena. For example, vertical strokes were exaggerated in space to account for smaller perceived shift magnitudes relative to horizontal strokes [43]. In some cases, curvature along stroke paths was also increased to enhance perceptibility. Initial stroke paths were designed to maximize detectability, through increases in stroke velocity and slight modifications to stroke geometry, within the first half-second of perception as

persistence of vision is established [44]. To objectively evaluate subtleties in design, each codex block was decomposed into space-time projections to calculation cumulative motion energy as a baseline for intra-codex comparisons. This process is detailed further in the results section.

3.1.1   Multi-aperture spatial arrangement
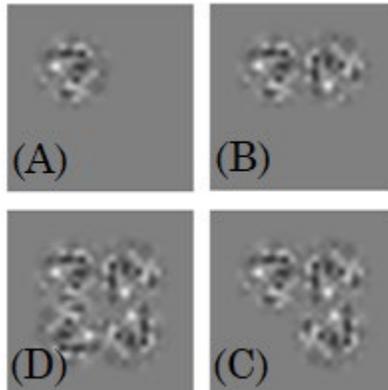
Figure 8. Motion-induced position shift aperture arrays for multi-stroke symbol representation. Multi-aperture symbols are reference by their stroke order: (A) 1st order, (B) 2nd order, (C) 3rd order, and (D) 4th order.

When delivering semantic cues, there immediately arises a need for expressing multi-stroke vector trajectories to communicate complex symbols within a moment arc of perception. To accomplish this, arrays of MIPS apertures were assembled, with the number of elements corresponding to the number of strokes required to impart minimum information necessary for symbol recognition.

What is important to note here is the displacement of additive carrier envelopes in relation to perceived structure. The most valuable asset of this approach is the <u>close onset, or compression, of physical loci within a confined visual environment</u>, i.e. retail space within an x-y coordinate plane or pixel pitch real estate of digital display technologies.[5]

---

[5] This becomes especially valuable in future application of the limited real estate of head mounted consumer technologies.

Apertures were activated consecutively in time and, when inactive, presented static grain. The spatial arrangement of arrays ensured each aperture was positioned relative to the previous stroke. For example, in the case of *A*, the first aperture contained the up-down stroke and the second aperture contained the right to left horizontal bisecting member. The second aperture is placed *above* and to the *left* of the first, positioning the second stroke in the correct location relative to the persistence of vision of where the first stroke ended.[6]

3.1.2   <u>Codex blocks</u>

The following plates on pages 30, 31, and 32 summarize stroke decomposition for twelve unique codex blocks used throughout experimentation. Strokes are color-coded according to order in sequence – 1st (green), 2nd (light blue), 3rd (dark blue), and 4th (orange). Note, sequential strokes with overlapping end-points and start-points are consolidated into one aperture (i.e. the second stroke of *B*). Separate apertures are only necessary when there exists a physical offset between strokes in perceptual space or, as in the petals on *flower*, when a continuous path is exceptionally long.

---

[6] In other words, for the letter *A*, as the first stroke turns downward after its peak, the second aperture is positioned such that its perceived location is correct relative to its offset in time from the persistence of vision of the end of the first stroke.
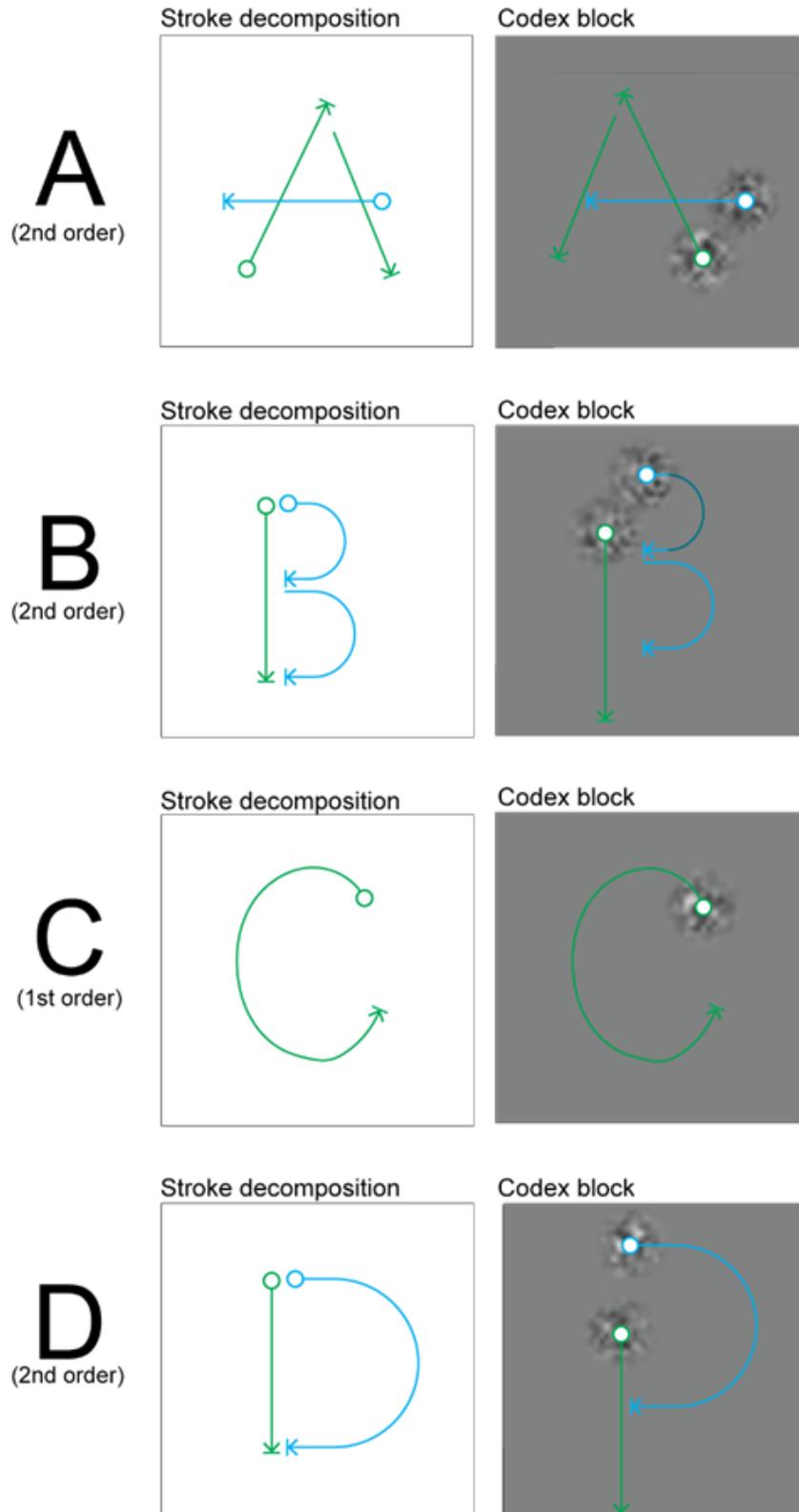
Figure 9. Symbol codex entries, { A, B, C, D }

3 | Contributions
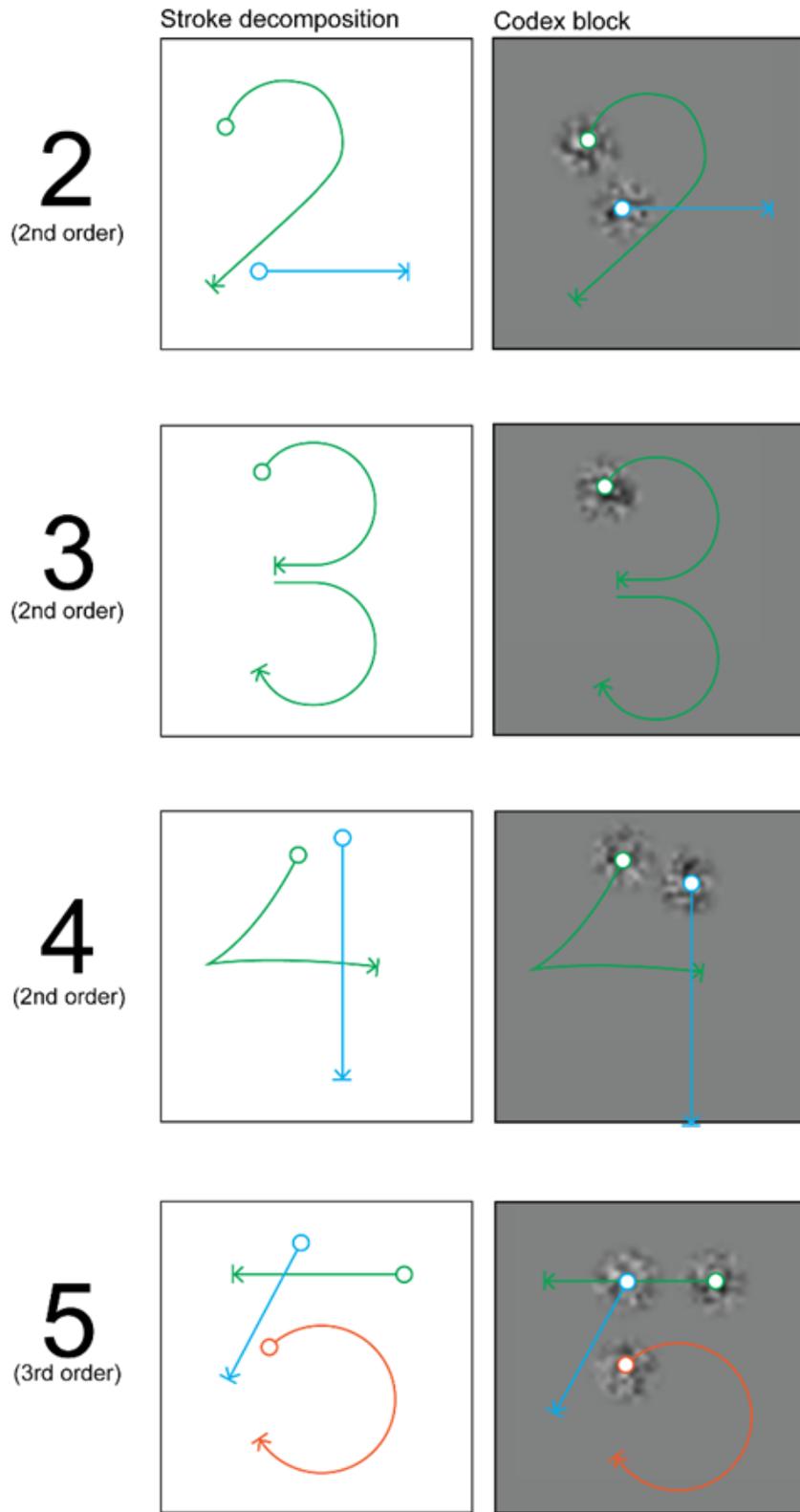
Figure 10. Symbol codex entries, { 2, 3, 4, 5 }
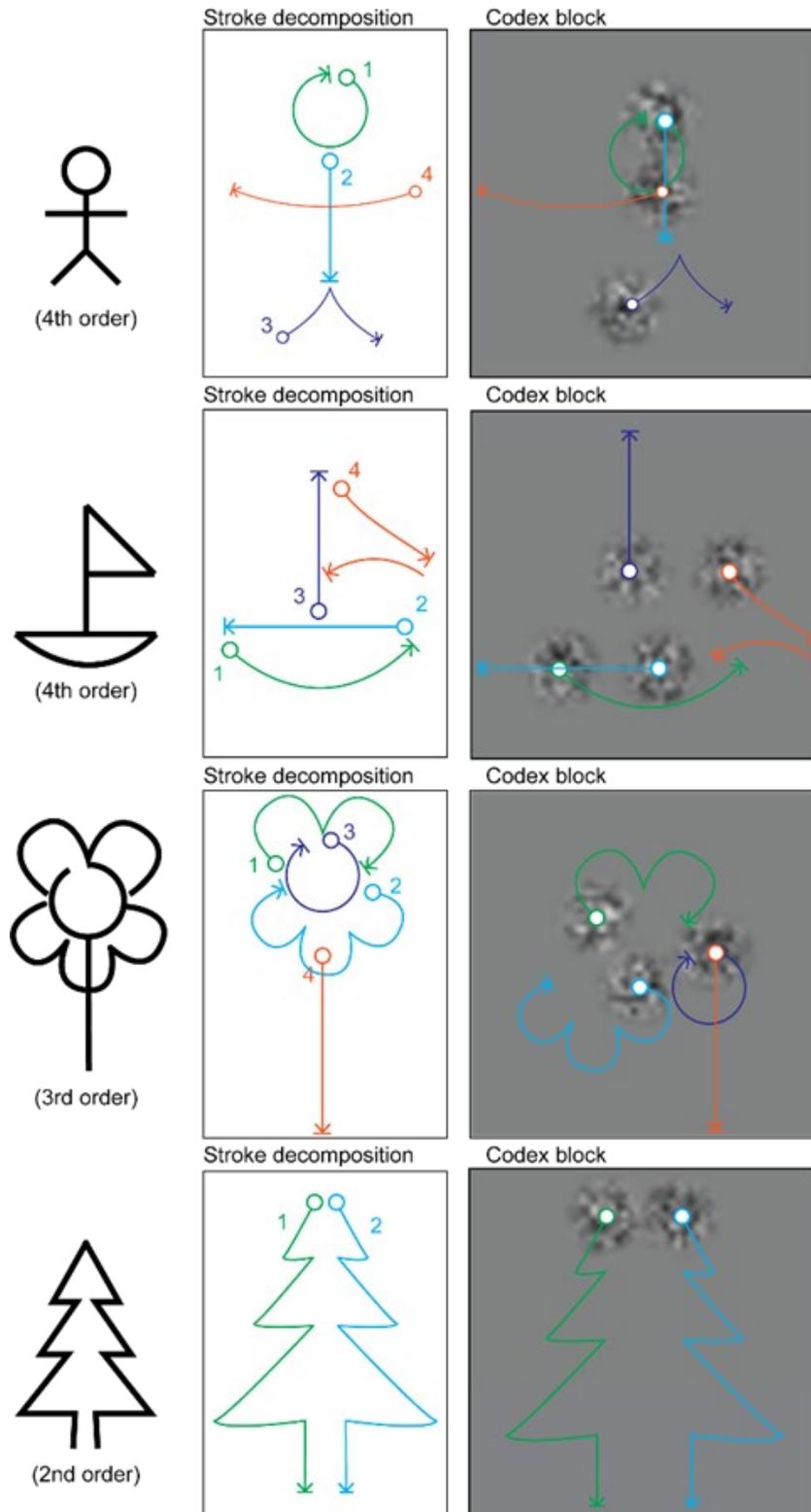
Figure 11. Symbol codex entries, common forms

Visual experience is a complex framework of perceptual interactions between motion and perception. As displays become larger, more immersive, and have the potential to extend beyond our central vision, we exploit several novel frameworks for augmented perception such as motion-induced position shift and slow speed bias, allowing us to engage the entirety of our visual field. Peripheral vision is capable of detecting complex motion cues. We are now learning how to computationally generate adaptations that allow you to read – via perceived object motion – from the edge of your vision. Using sensory plasticity as a medium to create new visual experiences allows us to challenge what we perceive as our extant realities.

3.2 Stimulus adaptation

One of the greatest challenges of adapting motion-modulated semantics to real-world environments is maintaining requisite conditions for effects to occur. Equiluminous kinetic edges are essential to induce perceived position shifts. The effect is somewhat tolerant to variance in luminance of the WGN patch when compared to the environment, but experiences a fall-off in magnitude with enough imbalance between them [30].
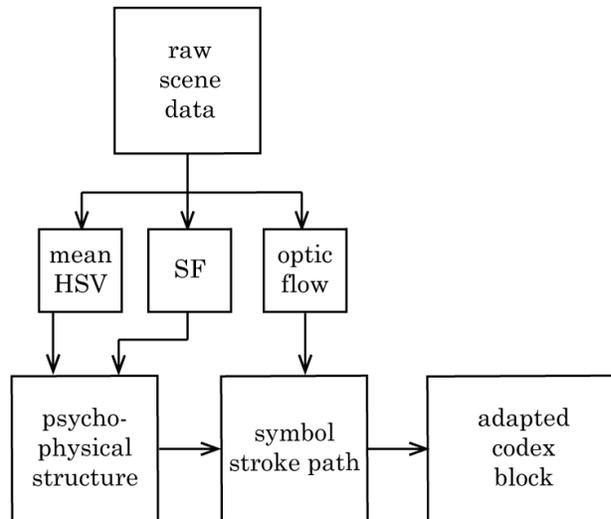
Figure 12. System block diagram of stimulus adaptation methodology for perceptual augmentation outcomes. HSV = Hue, Saturation,Value; SF = Spatial Frequency

This dissertation presents a two-stage peripheral stimulus generation approach. First, psychophysical structures are rendered and prepared to be animated behind an aperture along a stroke path defined by the symbol vector geometries. The second stage is a process of adapting peripheral stimuli to parameters from dynamic natural scenes for perceptual

augmentation in real-world environments. In this approach, raw scene data is parsed frame-by-frame into its chromatic, spatial frequency, and optic flow parameters. These outputs are applied to a raw psychophysical structure such as a sinusoidal grating or Gabor patch and its motion path geometry and adapted stimulus is then overlaid onto the environment and presented to the observer.
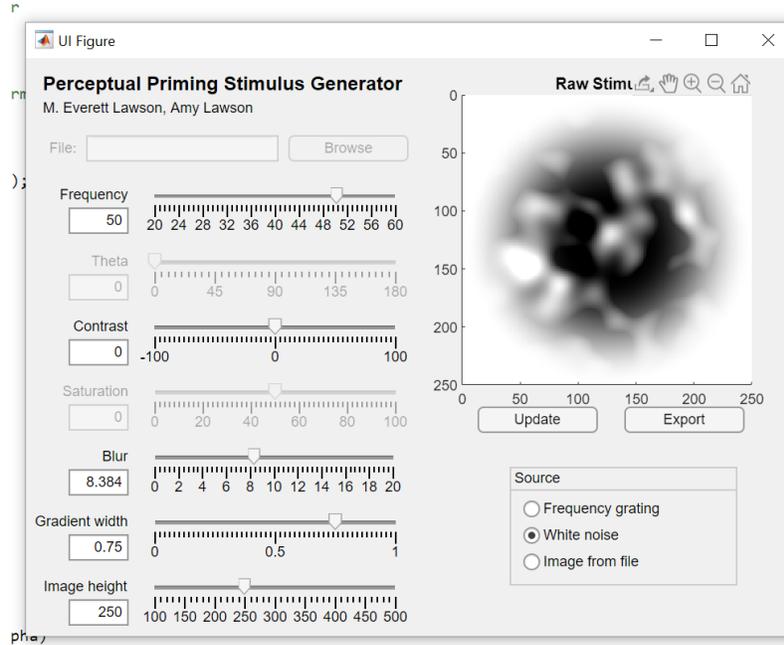
Figure 13. Program for generating raw psychophysical stimuli. Compatible source inputs include sinusoidal gratings, white Gaussian noise, and square ROIs from image files.

Generalized process for stimulus adaptation is detailed below:

1. Regions of interest (ROIs) are identified within the footage.

2. Scene parameters of hue, saturation, value, spatial frequency, and optical flow are calculated within ROIs

3. Codex block locations are identified within the simulated environment within the observer field of view.

4. Apertures are generated (either traditional Gaussian windows or scene-adaptive, non-uniform envelopes, measured in arc min)

5. Raw psychophysical carrier signals are generated.

6. Stroke paths for stimulus motion are generated.

7. Adapted codex blocks are generated from scene data and motion parameters and then integrated at defined carrier signal locations within the simulated environment.

3 | Contributions

Possible implementations of these methodologies would utilize powerful real-time video processing mechanisms to capture and parse environmental parameters. In the case of augmented driving experiences, increasingly sophisticated onboard processors developed for self-driving functions can be leveraged for this purpose. For future study design and experimentation, post-processing is practical for generating complete video sequences for presentation in a simulated environment.

Peripheral vision plays many roles in the manifestation of visual perception and, beyond information delivery, can be targeted in other new ways to affect perceptual outcomes. Vection, or perceived self-motion through visual stimulus alone, is heavily influenced by peripheral cues and creates strong illusory effects for the observer [45], [46]. As a tangible example, discordance between visual and vestibular sensory input can result in kinetosis, or motion sickness. There is an exciting opportunity to deploy peripheral stimuli to alter vection dynamically in real-world scenarios, not only for motion sickness mitigation but also to augment perceived driving experience[7]. In either case – for information delivery or for altering observer proprioception – an essential component of this systems approach is the contextual adaptation of peripheral stimuli, enabling application of these methodologies outside of controlled laboratory conditions.

3.2.1   Chromatic adaptation

In studies implemented throughout dissertation, chromatic adaptation was the primary focus for preparing codex blocks for dynamic visual environments. While spatial frequency and optical flow scene parameters are powerful tools in designing peripheral information stimuli, study design focused primarily on establishing first principles of far peripheral symbol recognition with minimal adaptation.

Sample adapted psychophysical structures are shown in Figure 14. Source psychophysical structures are second-order Gabor patches: windowed Gaussian white noise blurred with a

---

[7] Please see "Appendix A: Proprioceptive augmentation" for further descriptions of systems implemented over the course of this dissertation work to adaptively alter vection in dynamic visual environments.

2-D Gaussian smoothing kernel, σ = 5. Two potential methods of adaptation are highlighted here – overlay blending and 1:1 HSV adaptation. The former is a variation of element-wise multiplication of two source images resulting in linear interpolation of the top layer between the span of values in the base layer:

$$f(a,b) = \begin{cases} 2*a*b, & 0 \leq a < 0.5 \\ 1-2*(1-a)*(1-b), & 0.5 \leq a \leq 1 \end{cases}$$

Alternatively, HSV adaptation transforms the source patch RGB matrix to HSV space and updates hue and saturation layers to match mean values from environmental footage.
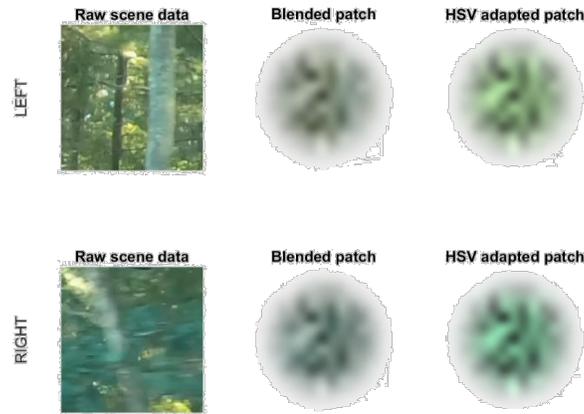
Figure 14. Sample adapted psychophysical structures demonstrating overlay blending (middle) and HSV adaptation (right). Source frame from the raw scene data used for adaptation is indicated on the left.

Overlay blending was used throughout the experimentation conducted to support this dissertation, favored for its subtler integration of stimuli within the environment. Future work would benefit the methodologies by quantifying relative outcomes of using either adaptation technique.

Figure 15. Single video frame from the high-load series experiment, automotive cockpit footage. The codex block, in this instance presented to the left visual field, is indicated by the red arrow

3 | Contributions
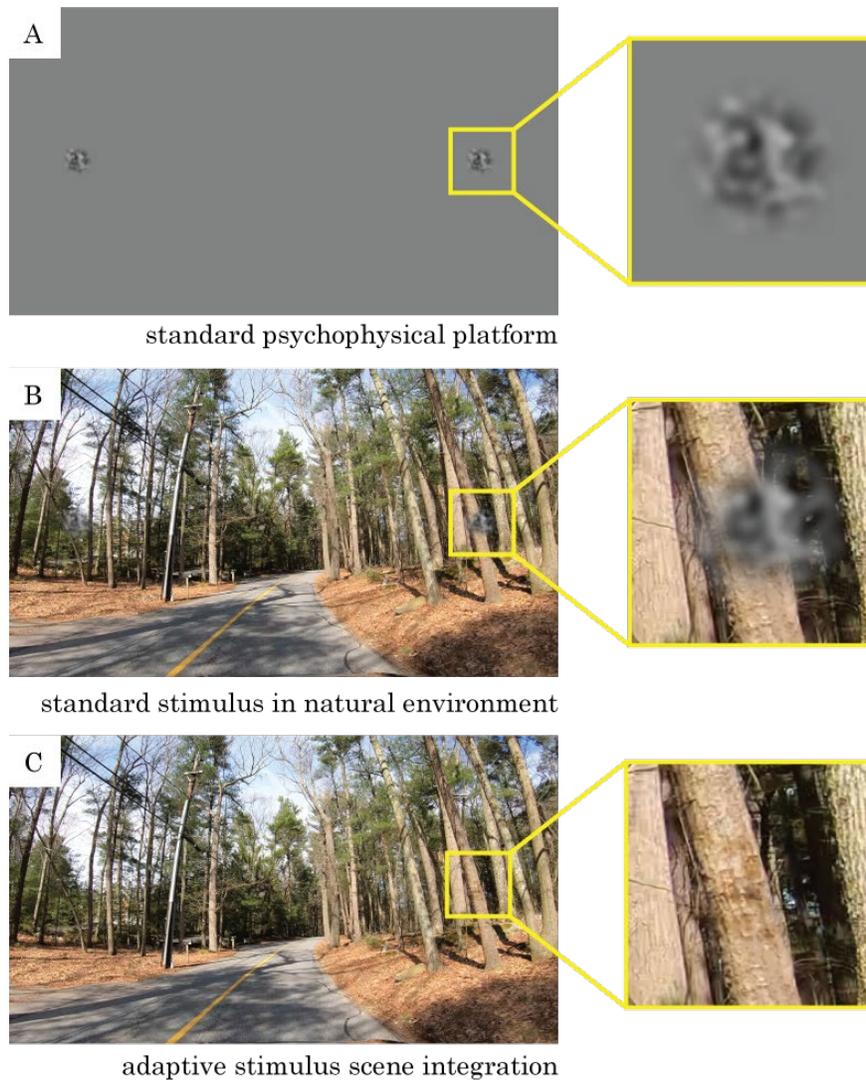
Figure 16. Windowed Gaussian noise comparison, (a) standard psychophysical stimulus and environment, (b) standard stimulus and natural environment, and (c) our method of stimulus overlaid with the natural environment

Perceptual shifts achievable with equiluminous kinetic edges, multidimensional or otherwise, are traditionally established and evaluated in a psychophysical setting (i.e. in a darkened room, viewing screens with controlled luminance ratios between the stimulus and the surround)[3]. The methods of adaptation and integration in dynamic environments presented here are not intended to diminish the efficacy or significance of psychophysical methods, which contribute powerful metrics to many fields, including vision research and cognitive psychology. Instead, the adaptation methodologies described in this dissertation are a motion toward broadening the scope of some of the most potentially impactful discoveries in vision research.

4. Related work

One of the greatest potentials inherent in activating fundamentals of vision science to create new media is the possibility of expanding or enhancing visual perception beyond its traditional scope or ability. This potential is the motivation behind some of the research most closely related to this dissertation.

4.1 Peripheral semantics

Many existing systems target peripheral vision as a gateway for information delivery to the user. Early successful implementations of peripheral cues in high-load tasks include the peripheral vision horizon display, first proposed by NASA in the 1960s [47] and later formalized to deliver orientation information to aircraft pilots [48]. Current applications have evolved to respond to the deluge of digital information available to human sensory modalities, in search of new ways to optimize presentation. Peripheral head-mounted displays have evolved to deliver information without obstructing the central field of view, though they traditionally require gaze deviation to convey complex concepts [49] Flicker, both highly perceptible and easily implemented via low-power LED arrays, has been integrated into several glasses-based designs for near-eye peripheral notification [50]–[55]. While a robust method in varied environments, flicker is fundamentally limited in the complexity of information it can convey. Beyond notification tasks, glasses-based peripheral augmentation has also been implemented for navigation or situational awareness for pedestrians [56], [57], or cyclists [58].

Luyten et. al. demonstrated a near-eye display for conveying symbols and shapes, both in static presentation and with applied rotation or translation [59]. The apparatus consisted of an acrylic frame worn as a pair of glasses which positioned two displays measuring 74 by 60 mm on either side of an observer's visual field. They found low detection rates for static symbols that improved when motion was applied to the stimulus or when interaction was included in the task. Further, they found that limiting the number of possible stimuli from which to discriminate increased recognition rates. While similar in scope and aim, the systems described in this dissertation require substantially less real-estate within the visual field to articulate: concepts such as *boat* or *flower* occupy less than 2 degrees. Further,

the method of generating peripheral stimuli by transforming characters and abstract symbols into perceived motion paths can be practically expanded to a broad range of information while still maintaining small form factors. In both this system and prior work, however, the number of possible symbols in the library or codex is a key determining factor in how quickly and with what confidence observers can perform accurate discrimination.

4.2 Automotive

The driving environment poses a significant opportunity to multiplex information delivery to vehicle operators. Systems targeting peripheral vision in this context focus primarily on delivering environmental notifications to the driver. Peripheral flicker stimuli have been deployed as components in advanced driver assistance systems (ADAS) [60], in automated driving systems for situational awareness [61], and via head-mounted displays that communicate directly with the vehicle [62]. Beyond flicker, ambient light has been implemented for both lane change alerts and navigation, using programmable LED strips to create apparent motion for notification [63], [64].

In one of the most relevant applications of motion-induced position shifts in this context, Kooi and Mosch demonstrated low contrast motion patterns overlaid on vehicle navigation displays, generating semantic motion cues to indicate navigational direction, notification, and urgent alerts [65]. The implemented stimulus occupied 9.52 degrees of the operator's visual field and was limited to a handful of coded messages, but was shown to effectively communicate information to the driver as well as alert to danger. While significantly smaller and thus likely less perceptible, the systems described in this dissertation can convey a much greater breadth of information.

In one of the first demonstrations of peripheral motion cues for augmenting vection in an automotive environment, Meschtscherjakov et. al. recently demonstrated a system for altering an observer's perceived velocity via apparent motion along a strip of LEDs within the vehicle [66]. While limited to the location of the LED strip within the cockpit, the system demonstrated that not only pixel speed but also temporal frequency and brightness of the light strip influenced perceived velocity. Systems for augmenting vection described briefly in this dissertation would utilize projection to coat the car interior in a *perceptual skin* resulting in a more global perception of motion stimulus with likely greater augmentation

magnitude. However, projection systems are susceptible to environmental light pollution while LED lighting remains highly detectible in varied environments. Ultimately, despite lighting challenges, a projection system could generate a far more dynamic and adaptable stimulus than a static LED array.

4.3 Virtual and augmented reality

Peripheral vision adopts the familiar role of guiding attention in many AR and VR applications (see [56], [67]). More often, the unique features of peripheral vision are activated to exclude them from central tasks, leveraging gaze detection to adaptively redistribute or reformat information within the display environment. Ohshima et. al. described methods for lowering detail in renderings of graphical objects that fall outside of fixation [68] and other studies have similarly tracked gaze to dynamically adapt information position and level of detail in AR/VR [69], [70].

Allocation and attentional focus of spatial resolution is a critical consideration in all near-eye display technologies. Early virtual reality systems were especially limited in visual field, but improvements in technology are now able to target far periphery. Sun and Varshney measured cortical scaling of static symbol detection at 10°, 20°, and 40° eccentricity in virtual reality environments, demonstrating methods for optimally scaling content for accurate detection [71]. Beyond this demonstration, however, little work has been done to activate peripheral vision as a dedicated channel for complex information delivery in augmented and virtual reality. The small visual form factor required to implement the codex blocks described in this dissertation makes these methodologies uniquely suited as a foundation for new media formats in this space.

5. Study methodology

The goal of human subject trials conducted throughout this research was to test the mechanics of fundamental building blocks of vision on the retina – to propagate a codex of stimuli with measurable impacts on human visual perception in dynamic environments. Study sequences drew from standard psychophysical experimental protocols to generate observer experiences that quantify various codex block parameters. During the study, participants performed simple discrimination tasks through self-reporting which were recorded against eye movements to indicate approximate visual field regions activated by our instrumentation and rendered media[8].

Key quantitative metrics yielded by experimentation were <u>symbol detection time</u> and <u>symbol detection accuracy</u> for each subject across varied visual environments.

5.1 Simulated environment

5.1.1 Design considerations

Evaluating adaptive cues in dynamic scenes requires a carefully controlled environment that is both parameterized and repeatable. Real, dynamic environments make poor study conditions for obvious reasons. Simulating these conditions, however, suffers from unavoidable limitations. The goal of the design of any simulated environment is to minimize these limitations as much as possible. Structure and capabilities of this environment necessarily dictate the content that is created for it, and for this reason refinement of the environment and capture of raw footage to propagate it happened concurrently and iteratively.

Quantifying the impact of stimuli, adapted to the natural scene and delivered to the visual periphery, has several key requirements which a simulated environment must meet:

---

[8] The stimuli used in these experiments were calibrated so as to not include moving patterns that are associated with risks of motion sickness or photosensitive epilepsy.

- *Visual field coverage*

- *Refresh rate:* sufficient to achieve smooth motion transitions (at least [x] Hz)

- *Brightness, saturation, and contrast:* within normal ranges for simulating outdoor conditions

- *Cost, financial:* budget for these early studies should accommodate quality components to reasonably ensure the integrity of results while remaining economical

- *Cost, temporal:* time required to set-up, calibrate, and operate hardware should be within reasonable limits

- *Compatibility:* hardware should be easily interfaced with necessary peripherals (gaze tracking, for example)

There are predominantly three methods for creating a simulated environment around an observer: using digital screens, projections, or virtual reality hardware. All three of these methods, in the right configurations, can meet the requirements above. The selection process is summarized in the table below: each option is given a score for each metric, which is totaled according to assigned weight for each metric.

Table 2. Comparison between methods for simulating a dynamic environment for observation.

Requirement	(weight)	Digital screen	Projection	VR
VF coverage	0.25	3	4	2
Refresh rate	0.1	5	5	5
B,S,C	0.1	5	4	5
Cost, (financial)	0.1	2	3	4
Cost (temporal)	0.25	4	2	2
Compatibility	0.2	5	5	3
TOTAL	1.00	3.95	3.7	3

In designing an environment for study, the two most significant considerations were complexity and visual field coverage. Digital screen solutions involve the simplest implementation, while projection requires the construction of a curved surface with geometric reform and VR requires content generation in unfamiliar formats. The digital screen solution, while not configurable to cover very large areas of visual field within a reasonable budget, proved simplest and most compatible for this use case while still meeting baseline requirements.

5.1.2 Observer geometry
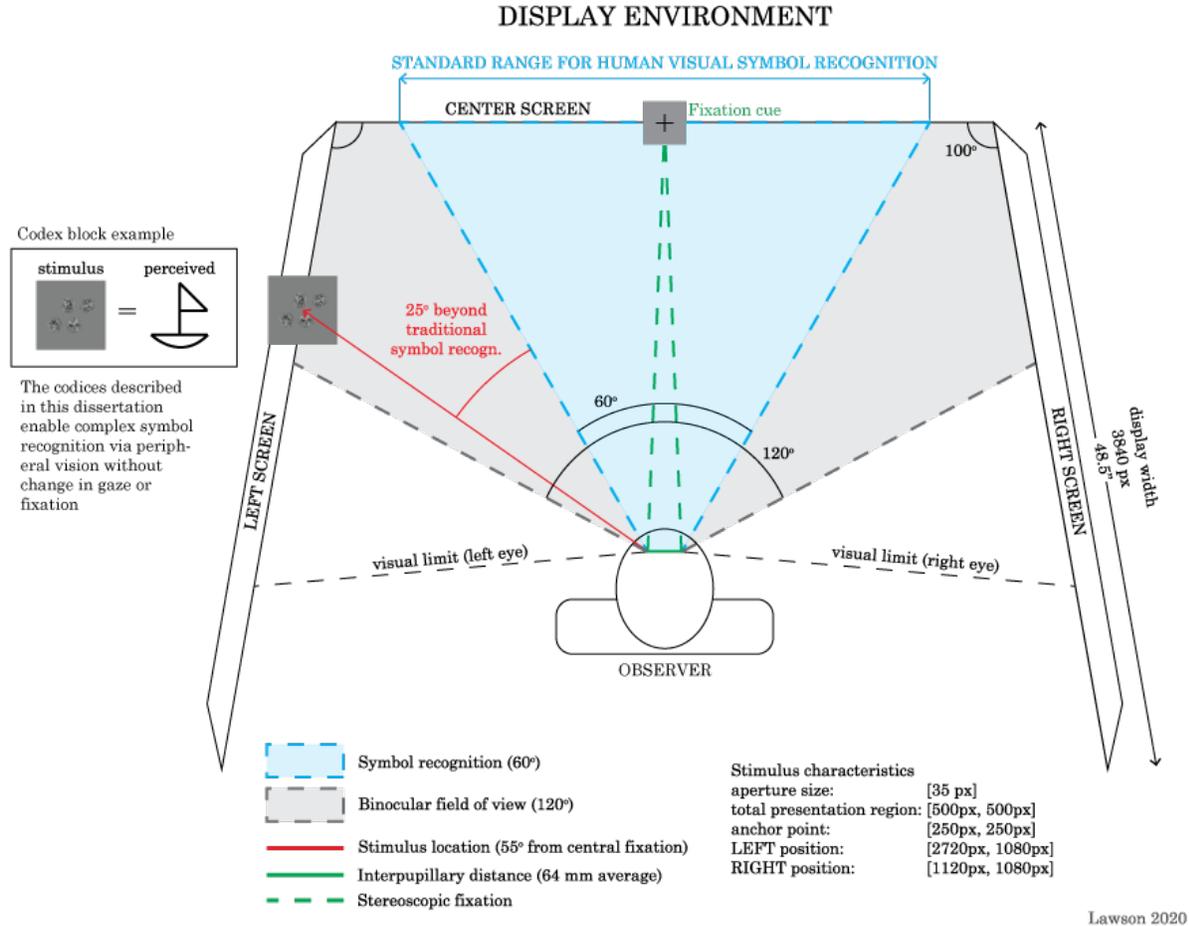
Figure 17. Ray diagram of the simulated environment

High-resolution LCD displays surrounded the observer: the center parallel to the observer's coronal plane and the left and right displays abutting, at each seam forming a 100° angle. The observer was seated 31.5 inches from the center screen such that each pixel subtended 1.10 arcmin of the visual field with no restraining mechanisms to fix head or body motion. Three displays were mosaiced using an NVIDIA Quadro RTX 4000 graphics card, driver version 431.02. Gaze tracking was captured using a Pupil Core Headset and calibrated and recorded using Pupil Capture software v1.16 to evaluate RTs as relative to gross motor and perceptual response times. Observers reported threshold detections using a standard US layout, wired keyboard input device (Amazonbasics model KU-0833), and inputs were timecoded using MATLAB 2019b.

5.1.3   Gaze tracking

For gaze tracking, an outfitted Pupil Core Headset was calibrated to each individual before beginning each study sequence. Sample calibration results are summarized in Table 3 below.

Table 3. Sample calibration results for Pupil Core Headset gaze tracker

Metric	Description	Obtained value (degrees)	Published value in ideal conditions [72] (degrees)
Angular accuracy	Average angular offset (in degrees of visual angle) between fixation locations and corresponding locations of fixation targets	1.46°	0.60°
Angular precision	Root mean square of angular distance between successive samples	0.05°	0.08°

5.1.4   Footage capture and stimulus preparation

Raw footage from natural environments was captured with a custom 12 DOF rig consisting of three GoPro HERO7 cameras, each recording 4K, 60 FPS with 69.5° vertical and 118.2° horizontal FOV and encoding using H.265 HEVC codec.

Figure 18. Vehicular system for raw footage capture and carrier signal delivery: (a) the three-view camera mount, (b) the projection system for delivering carrier signal cues and environmental overlays, (c) sample calibration pattern for mapping the vehicle interior

Source videos were captured from two separate environments: first, from a pedestrian point of view (POV) while walking and fixating on a handheld device, second, from a pedestrian POV while walking and facing forward, and third from the driver's POV in an automotive cockpit. Pedestrian footage was captured from a single video source while automotive footage utilized the three-sensor array. Footage was stabilized and aligned for display in the simulated environment using Adobe After Effects v17.0.2.

Figure 19. Registration marks for post-processing and stimulus positioning in three-view automotive footage. Full-field overlay shown here demonstrates an example of a projected architecture in the interior of a real-world environment (target: vection augmentation).

MATLAB 2019b was used to process footage for scene parameters, generate raw psychophysical structures, and adapt them to scene parameters. Structures were then animated according to motion data from the codex in Adobe After Effects. Adapted, animated stimuli and raw footage were integrated using Adobe Premiere Pro v14.0.1. Completed study sequences were displayed on three 65" 4K UHD displays (Samsung UN65RU8000FXZA) at 120Hz refresh rate.

5.2 Procedures

Table 4. Summary of series parameters for conducted experimentation.

Series	Subjects	Trials/ subject	Codex blocks/ trial	Stimulus repetitions/ codex block
Control	11	10	24	5
Low-load	2	10	24	5
High-load (shore drive pedestrian footage)	2	1	36	2 (for alphanumeric), 1 (for abstract)
High-load (town center pedestrian footage)	2	2	45	2 (for alphanumeric), 1 (for abstract)
High-load (automotive cockpit footage)	5	2	18	2 (for alphanumeric), 1 (for abstract)

Studies were designed to address two primary research goals: first, to establish perceptibility, as measured by recognition accuracy and speed, of the codex in standard psychophysical environments, and second, to evaluate how perceptibility changes with increasing scene complexity in the central visual field. The series, summarized in Table 4, first presented codex blocks with a standard 50 gray background and static central fixation ("control"). Next, in the low-load series, each codex block was evaluated with the same dynamic natural scene presented centrally, to normalize the effects of increasing scene complexity on codex

perceptibility. Finally, three high-load environments were tested, in which codex blocks were presented in sequence over continuous video clips, with no normalization for variations in the scene.

All participants conducted the studies binocularly in a dimly lit room. Before the start of each study sequence, gaze tracking was calibrated for each participant, and metrics for angular accuracy and precision were recorded. Eleven subjects having normal or corrected-to-normal vision participated in the series described below. Exempt status was obtained from the MIT Committee on the Use of Humans as Experimental Subjects for minimal risk research (Exempt ID E-1866).

5.2.1   Introducing symbols

Prior to starting the control series, each participant viewed an introductory sequence to become familiarized with the 12 codex blocks used throughout the study. The following three steps were repeated for each codex block:

1. The codex block was presented to the left and then right peripheral visual fields.
2. Its identifying symbol was presented to the central visual field.
3. The codex block was presented again to the left and then the right peripheral visual fields.

The 12 codex blocks and their identifying symbols are summarized in the table below.

Table 5. Summary of codex blocks and their representative symbols.

Symbol ID	Representation
A	A
B	B
C	C
D	D
2	2
3	3
4	4

5 **5**

Boat

Flower

Stick figure

Tree
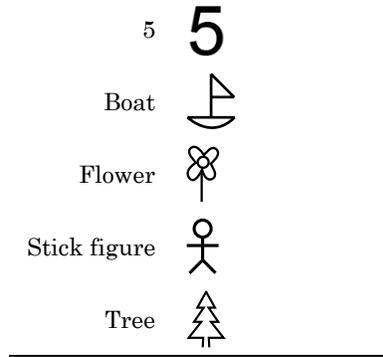
5.2.2   Control series (standard psychophysical environment)
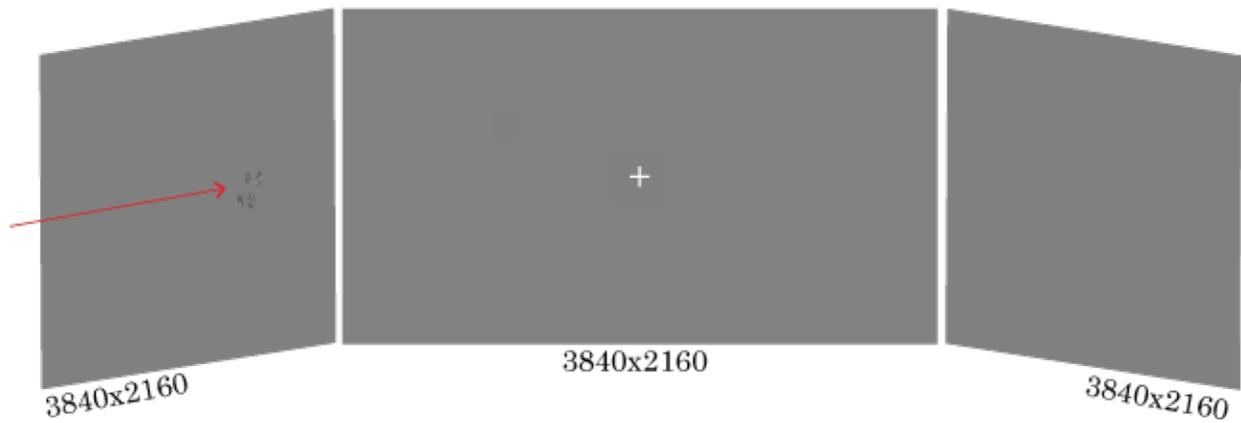
Control series summary			
Subjects	**Trials/subject**	**Codex blocks/trial**	**Stimulus repetitions/codex block**
11	10	24	5

Figure 20. Single video frame from the control series and data summary. This environment is maintained as constant throughout the series. The codex block, in this instance presented to the left visual field, is indicated by the red arrow.

11 subjects completed 10 trials each in the control series. Each trial consisted of a randomized sequence of 24 codex blocks. The randomized sequence drew from 12 unique symbols presented to either left or right visual field (totaling 24 instances within a trial). A standard psychophysical environment was maintained throughout the control series. A central fixation cue was presented on a 50 gray background while a random codex block was presented peripherally (55° off-axis) in either right or left visual fields. Each codex block consisted of 5 repetitions of the motion stimulus.

Subjects were instructed to press a button as soon as they recognized the symbol. Once the subject indicated detection, the stimulus was removed, and the subject was instructed to record their observation on the provided sheet.

5.2.3   Low-load series (repeating pedestrian POV environment)

The low-load environment consisted of the same dynamic natural scene presented to the observer's visual field for the evaluation of each codex block, to normalize any outcomes on perceptibility.

Low-load series summary			
Subjects	**Trials/subject**	**Codex blocks/trial**	**Stimulus repetitions/codex block**
2	10	24	5

Figure 21. Single video frame from the low-load series and data summary. The same video footage is repeated for each codex block in the randomized sequence. The codex block, in this instance presented to the right visual field, is indicated by the red arrow.

Two subjects from the control series group continued to complete the low-load series, 10 trials each. Each trial consisted of a randomized sequence of 24 codex blocks, drawn from the same 12 codex blocks from the control series, presented either to the left or right visual fields. Subjects were presented with video footage from a pedestrian point of view (POV) with no static fixation cue. The same footage was repeated for each of the 24 codex blocks in a trial. Five seconds elapsed between the start of the footage and the onset of the stimulus. Codex blocks were overlaid peripherally (55° off-axis) to either right or left visual fields. Each codex block consisted of 5 repetitions of the motion stimulus. Subjects were instructed to press a button as soon as they recognized the symbol. Once the subject indicated detection, the stimulus was removed, and the subject was instructed to record their observation on the provided sheet.

Throughout dissertation, a first-generation prototype was developed for delivering far peripheral complex information without foveal occlusion or interrupting medium between observer and environment (Figure 22). For purposes of subject safety and per the conditions

of COUHES approval, evaluation of this prototype was not within the scope of these foundational experiments. Given appropriate safeguards for navigating external environments, second and third-generation prototypes will be evaluated in real-world scenarios.
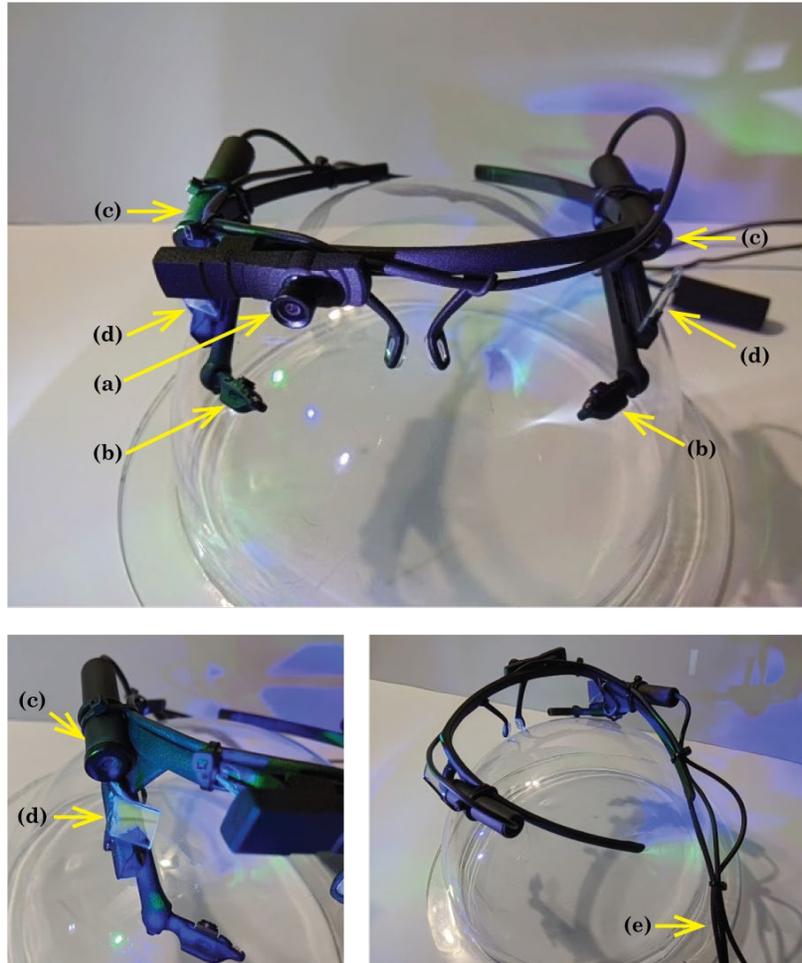
Figure 22. System, adapted from Pupil Core headset, for stimulus delivery beyond experimental settings. Functional elements: (a) 1080p30Hz "world" camera (Pupil Core architecture), (b) IR illumination and sensor for gaze capture (Pupil Core architecture), (c) pico projector for adapted stimulus delivery, (d) dichroic beam splitter projection surface positioned within observer peripheral visual regions, and (e) data and power cables.

5.2.4   High-load series (non-repeating dynamic POV environment)

The high-load series consisted of three separate trials drawing from three unique source footage environments: shore drive pedestrian, town center pedestrian, and automotive cockpit. Codex block presentation continued uninterrupted throughout the footage, in contrast to the repeated footage used in the low-load case. In each trial, the randomized sequence of codex blocks was overlaid 55° off-axis onto a continuous, non-repeating video with

no central fixation cue. Codex blocks were clustered into groups of three, with two seconds elapsing between each block within a group, and five seconds elapsing between each group in the sequence. The number of motion stimulus repetitions was reduced in the high-load case: symbols consisting of fewer than four strokes were displayed twice, while symbols consisting of four strokes were displayed once.

Subjects were instructed to press a button as soon as they recognized the symbol and then verbally announce symbol identification. Video played continuously with no pause for reporting, regardless of the accuracy of observer detection.

The two subjects who completed control and low-load series also completed the entire high-load series (shore drive pedestrian, town center pedestrian, and automotive cockpit). Three subjects from the control series group also completed the automotive cockpit high-load series.

High-load series summary (shore drive pedestrian footage)			
Subjects	Trials/subject	Codex blocks/trial	Stimulus repetitions/codex block
2	1	36	2 (for alphanumeric), 1 (for abstract)

Figure 23. Single video frame from the high-load series, shore drive footage and data summary. The 36 codex blocks are overlaid in randomized sequence onto the continuous video. The codex block, in this instance presented to the right visual field, is indicated by the red arrow.

5 | Study methodology

High-load series summary (town center pedestrian footage)			
Subjects	**Trials/subject**	**Codex blocks/trial**	**Stimulus repetitions/codex block**
2	2	45	2 (for alphanumeric), 1 (for abstract)

Figure 24. Single video frame from the high-load series, town center footage and data summary. The 45 codex blocks are overlaid in randomized sequence onto the continuous video. The codex block, in this instance presented to the left visual field, is indicated by the red arrow

High-load series summary (automotive cockpit footage)			
Subjects	**Trials/subject**	**Codex blocks/trial**	**Stimulus repetitions/codex block**
5	2	18	2 (for alphanumeric), 1 (for abstract)

Figure 25. Single video frame from the high-load series, automotive cockpit footage and data summary. The 18 codex blocks are overlaid in randomized sequence onto the continuous video. The codex block, in this instance presented to the left visual field, is indicated by the red arrow

6. Results

6.1 Gaze diversions

Three-dimensional gaze position coordinates were captured for all subjects throughout experimentation. Gaze data for control series was synchronized to stimulus onsets using MATLAB 2019b and then parsed into segments corresponding to each codex block presentation. A simple fixation detector was implemented to (1) filter gaze velocity for artifacts exceeding physiological limits (900 deg/sec)[73], (2) identify fixations with a minimum duration of 80 milliseconds, and (3) detect instances of gaze diversion in excess of 30 degrees from central stereoscopic fixation. Symbol presentations with significant gaze diversion were excluded from final results.

6.2 Symbol motion energy analysis (MEA)

To objectively evaluate the many varied stimuli used throughout the studies, motion energy was calculated for each adapted carrier signal using a multi-stage analysis described by Lu and Sperling [74]. First, each codex block was decomposed into two projections on the x-t and y-t planes. Fourier power spectra were computed for each projection to isolate harmonics. Total motion energy is given by summing the contributions of each quadrant of the power spectrum:

$$ME = (E_1 + E_3) - (E_2 + E_3)$$

For E_i, where i denotes the quadrant of the power spectrum when plotted on fx-ft and fy-ft planes. Motion energies of each projection, i.e. Ex (x-projection motion energy) and Ey (y-projection motion energy), form components of a vector in space whose magnitude yields absolute motion energy E_{total}.

$$E_{total} = \sqrt{E_x^2 + E_y^2}$$

The sign of components E_x and E_y denotes the cumulative direction of motion along each projection. Positive E_x indicates motion from left to right while negative E_y indicates motion from right to left; positive E_y indicates motion from top to bottom while negative E_y from bottom to top.
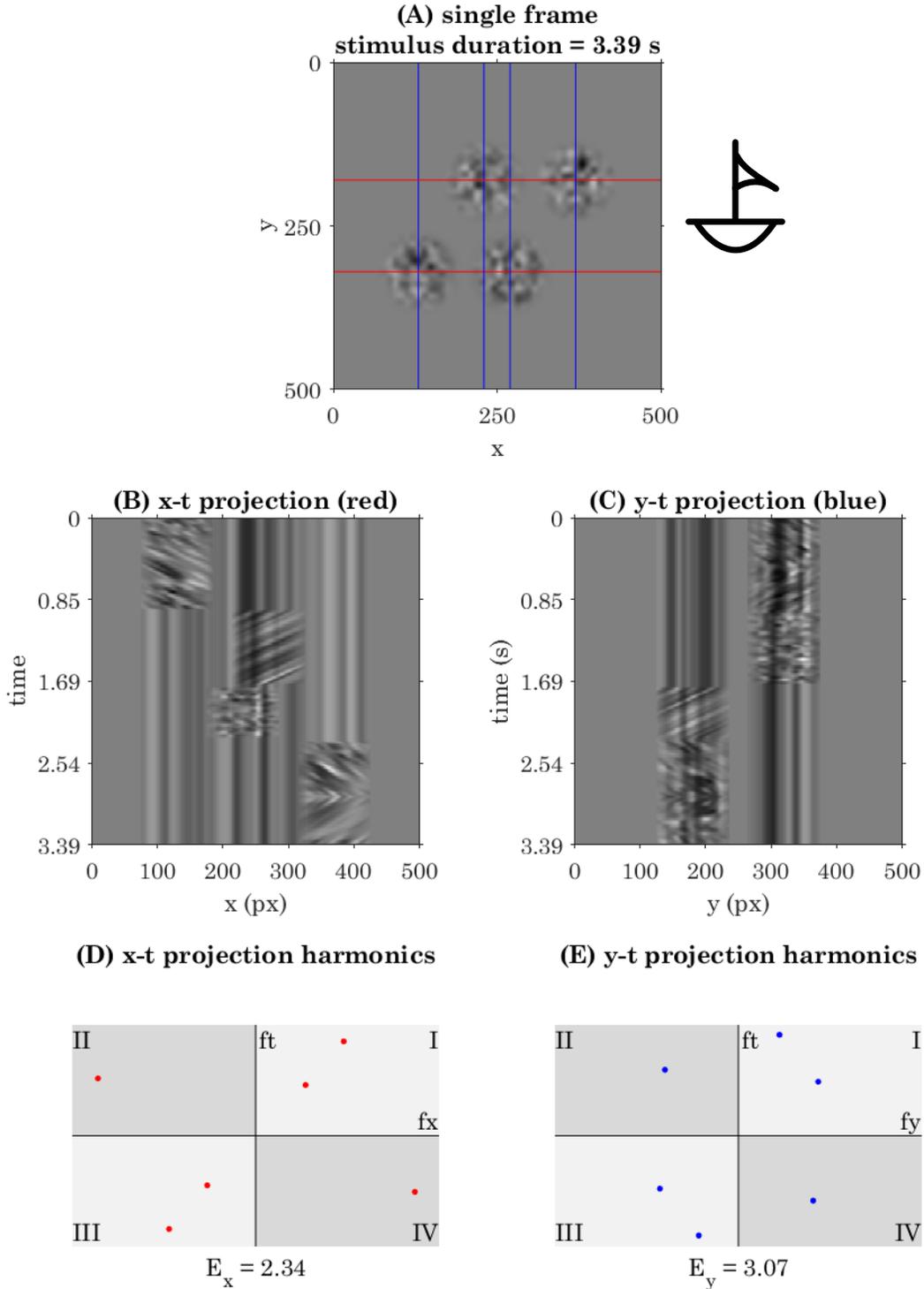
Figure 26. **Motion energy analysis for BOAT codex block. (A)** Single frame excerpt of stimulus video with cross-section annotations indicating locations of x-t projections (red) and y-t projections (blue). **(B, C)** Projections in x-t and y-t space at cross-sections indicated in (A). Each column represents a fixed aperture in space. The absence of motion in a given aperture is represented by repeating vertical pixels. (D, E) Fourier analysis of motion energy in fx-ft and fy-ft planes with aggregate motion energy value indicated below. Harmonics from the associated projection are indicated by corresponding color (x – red, y – blue).
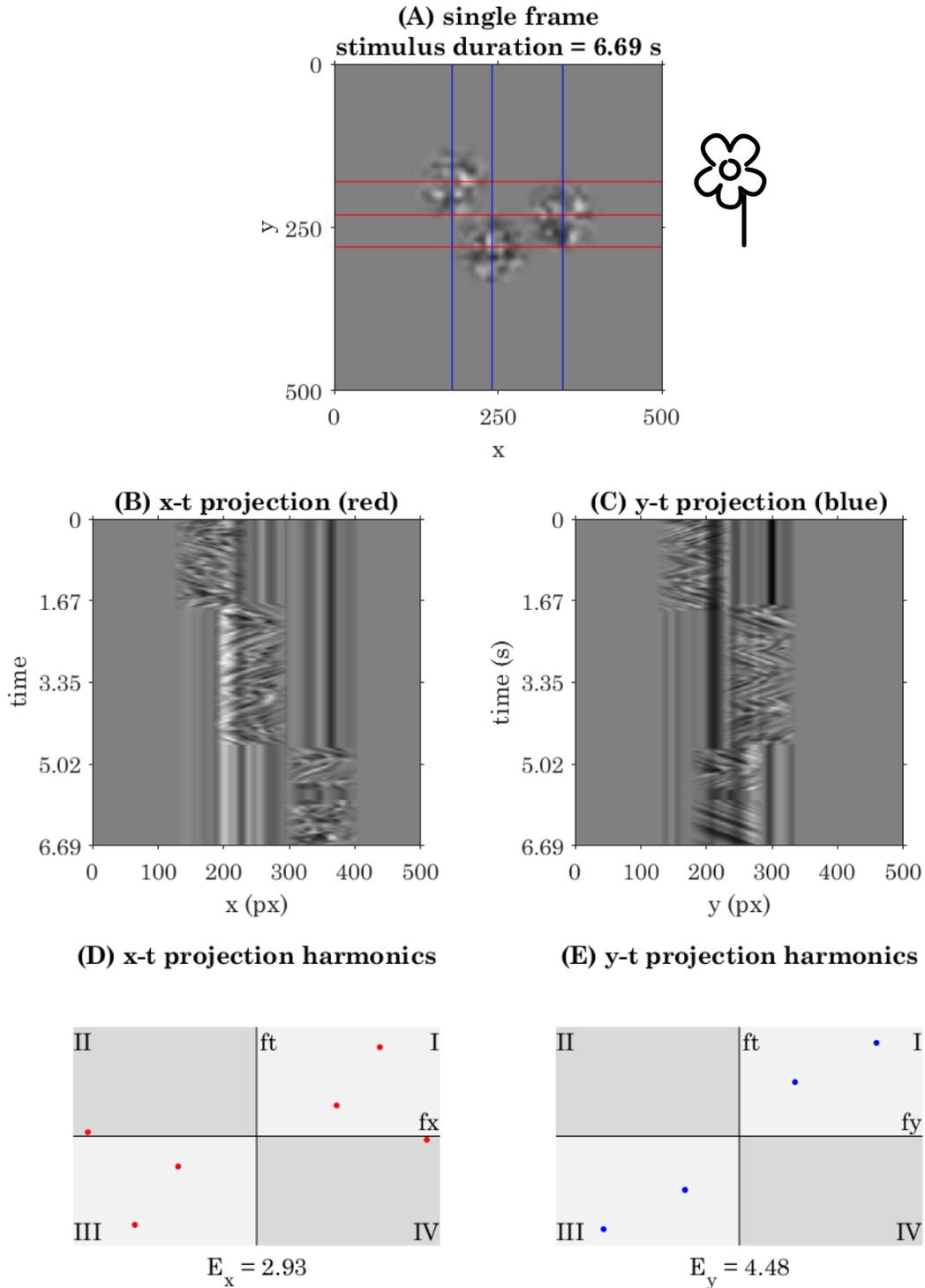
Figure 27. **Motion energy analysis for FLOWER codex block. (A)** Single frame excerpt of stimulus video with cross-section annotations indicating locations of x-t projections (red) and y-t projections (blue). **(B, C)** Projections in x-t and y-t space at cross-sections indicated in (A). Each column represents a fixed aperture in space. The absence of motion in a given aperture is represented by repeating vertical pixels. (D, E) Fourier analysis of motion energy in fx-ft and fy-ft planes with aggregate motion energy value indicated below. Harmonics from the associated projection are indicated by corresponding color (x – red, y – blue).

6 | Results
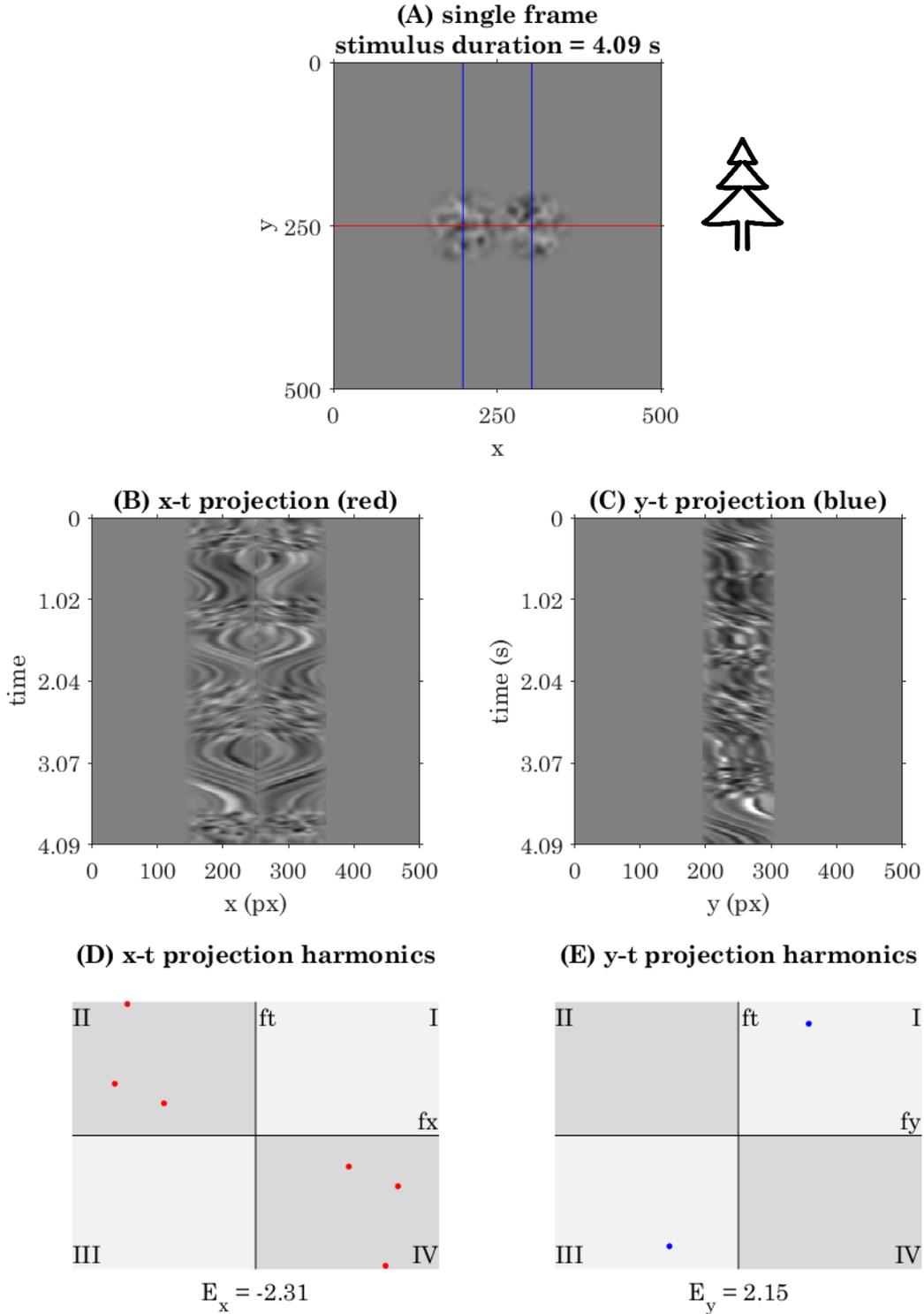
Figure 28. **Motion energy analysis for TREE codex block. (A)** Single frame excerpt of stimulus video with cross-section annotations indicating locations of x-t projections (red) and y-t projections (blue). **(B, C)** Projections in x-t and y-t space at cross-sections indicated in (A). Each column represents a fixed aperture in space. The absence of motion in a given aperture is represented by repeating vertical pixels. (D, E) Fourier analysis of motion energy in fx-ft and fy-ft planes with aggregate motion energy value indicated below. Harmonics from the associated projection are indicated by corresponding color (x – red, y – blue).

Plates on pages 53, 54, and 55 represent examples of codex blocks decomposed into space-time projections to yield objective measurements of motion energy[9]. Motion path geometries for each symbol are indicated to the left of (A) "Single frame". X-t and y-t projections in Figure 26 (B) and (C) are the result of superimposing projections from each aperture into a single image. Combining projections allows us to evaluate total motion energy from all apertures throughout the entire codex block duration. Cross-sections in Figure 26(A) indicate the location of each projection (x-t in red, y-t in blue). On-off timing of each aperture is evident in x-t and y-t projections by repeated rows of pixels indicating the absence of motion.

As described in the chapter on contributions, each aperture partially occludes an image consisting of white Gaussian noise – this WGN image is translated along a stroke path defined by the symbol. The impact of non-uniformity in the WGN image on motion energy is evidenced in symbol *TREE*. This codex block displays two synchronous, symmetric strokes, each forming either side of the tree symbol. If this codex block were designed with identical images behind each aperture (such as phase-matched spatial frequency gratings or the same WGN mirrored about its vertical axis), motion energy analysis along the x-projection would sum to zero. In this case, non-uniformities in the grain generate resultant motion artifacts which skew energy component E_x to -2.31, indicating the presence of motion from right to left. For comparison, other codex blocks composed of single apertures and approximately neutral motion in x-projection such as *3* and *C* exhibit E_x values at or near zero.

The following figure on (page 57) represents the distribution of motion energy results for all symbols in the codex. Horizontal and vertical axes represent x- and y-projection motion energy respectively. Velocity vectors in the E_x-E_y plane represent orientation and magnitude of resultant motion at each coordinate in space, given by $(u, v) = (x, -y)$. Characteristics of each codex block are summarized in the table below with absolute motion energy E_{total}.

---

[9] Motion energy analysis results for the entire codex can be found in "Appendix C: Supplemental results".
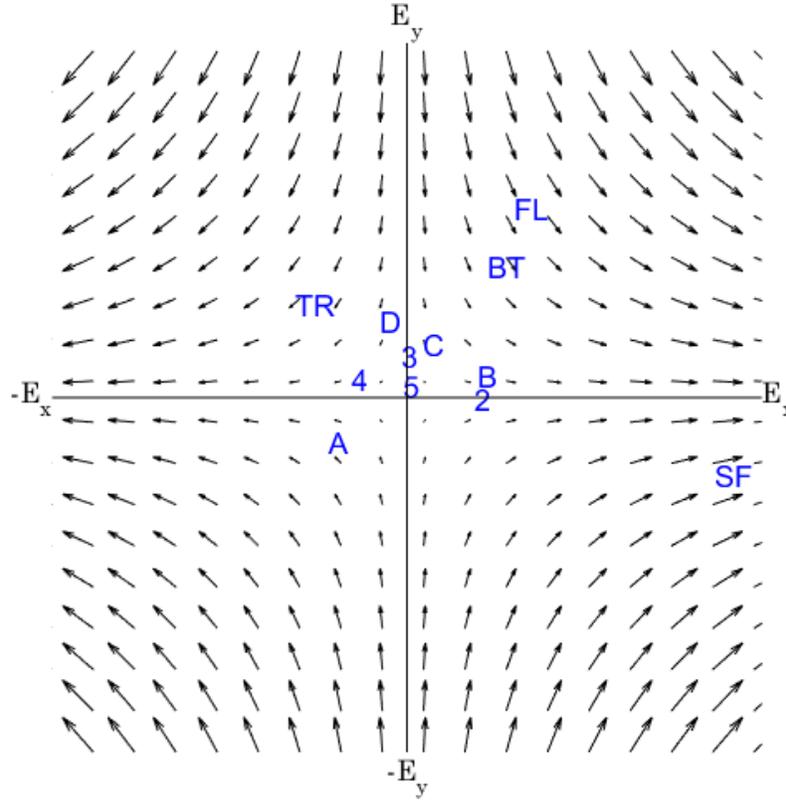
Symbol abbreviations

BT = ⏚ FL = ✿ SF = 人 TR = 🌲

symbol	apertures	duration (s)	E_x	E_y	E_{total}
2	2	3.050	1.737	-0.151	1.744
3	1	3.184	-0.043	0.896	0.897
4	2	2.784	-1.245	0.350	1.293
5	3	3.517	-0.019	0.578	0.578
A	2	3.517	-1.753	-1.181	2.114
B	2	4.250	1.866	0.433	1.916
C	1	2.484	0.547	1.182	1.302
D	2	3.100	-0.468	1.745	1.807
BOAT	4	4.868	2.337	3.065	3.854
FLOWER	3	8.167	2.929	4.479	5.352
STICKFIGURE	3	6.534	7.814	-2.002	8.066
TREE	2	5.567	-2.311	2.147	3.154

Figure 29. Symbol motion energy analysis, represented in E_x-E_y space, for 12 unique symbols with varying durations and stroke paths. Motion energy sign indicates aggregate symbol motion direction: positive E_x indicates motion from left to right, negative E_x indicates motion from right to left; positive E_y indicates motion from top to bottom, negative E_y indicates motion from bottom to top.

6.3 Detection time by symbol
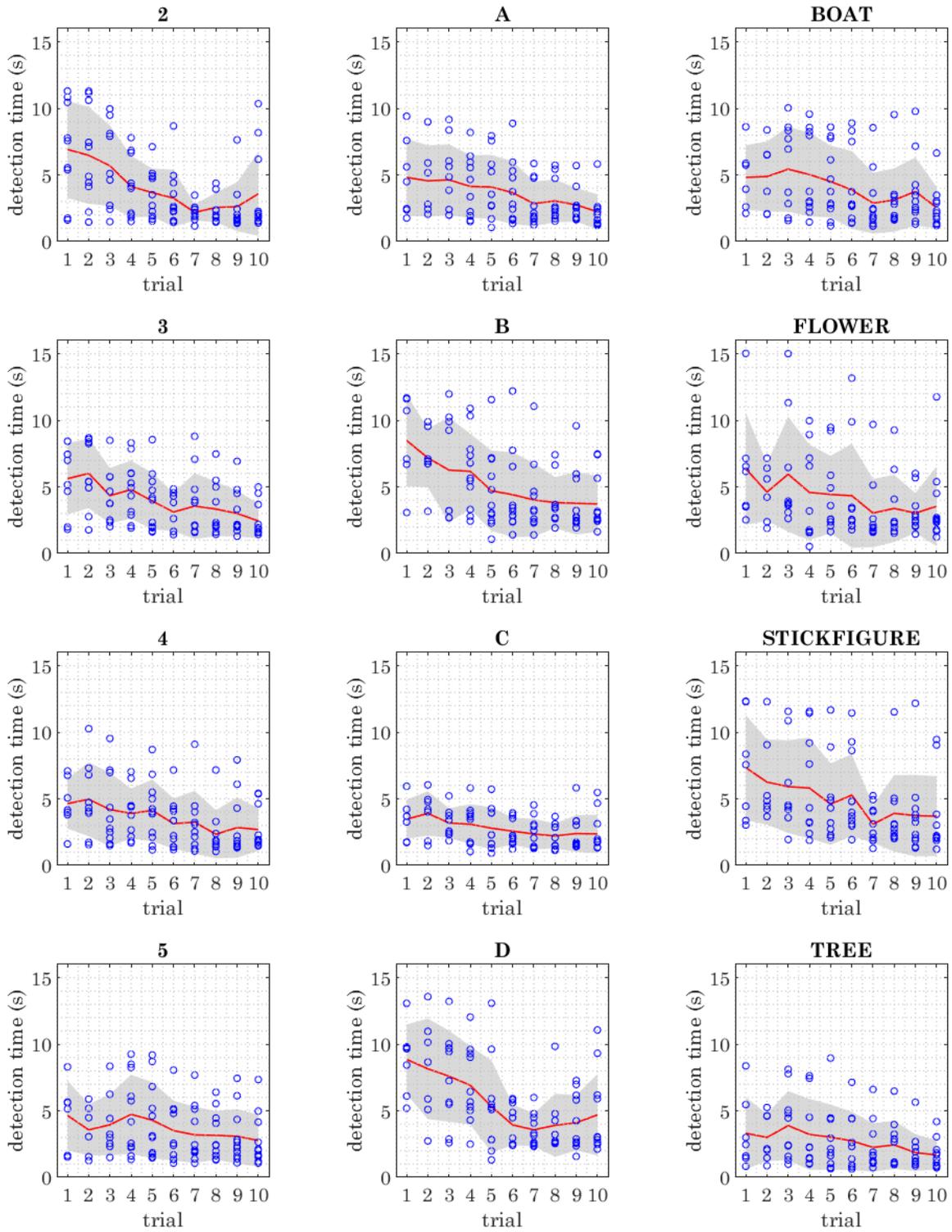
Figure 30. Symbol speed of detection across successive trials in control series (n=11) for all codex blocks. Mean detection time is indicated by the red line; raw detection times are indicated by the unfilled blue circles. Detection time span of one standard deviation from the mean $(\mu \pm \sigma)$ is indicated by the shaded region.

Speed of detection for each symbol over successive trials is represented in Figure 30. Raw detection times for all subjects were collected and evaluated for accuracy. <u>Detection responses corresponding to either incorrect or absent symbol identification were removed</u>. Mean response time for each trial is indicated in red with a shaded area representing the span of responses (trial mean ± one standard deviation).

6.4 Detection accuracy

Correct symbol (high-confidence) confirmation was recorded via button press within the window of stimulus presentation coupled with accurate symbol ID reporting on the provided sheet (in addition to recorded verbal response) and weighed against gaze threshold limitations. Instances of absent or late detection button presses or inaccurate symbol ID reporting were <u>logged as incorrect</u>. In the control series, conducted in a standard psychophysical environment, accuracy rates were consistently high, with little variance across subjects (Table 6).

Table 6. Mean symbol detection accuracy by trial, control series (n = 11).

trial	symbols/trial	mean (μ)	variance (σ^2)
	Symbol detection accuracy (control series, n = 11)		
1	24	0.875	2.82E-02
2	24	0.967	2.21E-03
3	24	0.985	7.84E-04
4	24	0.977	1.16E-03
5	24	0.979	1.23E-03
6	24	0.958	4.62E-03
7	24	0.975	2.03E-03
8	24	0.981	1.37E-03
9	24	1.000	0
10	24	0.975	4.84E-04

High rates of symbol detection accuracy were maintained with increased environmental distractors in low-load pedestrian POV series. A comparison between symbol detection accuracies for Subjects {1,2} across both control and low-load series (Table 7) shows that high

accuracy rates achieved throughout control trials were carried over into low-load series. High-load series accuracy rates, while initially decreasing to levels comparable to the initial trial in control series, rebound to >90% accuracy for the remaining high-load trials.

Table 7. Mean symbol detection accuracy by trial, control, low-load, and high-load series (n=2). †pedestrian POV environment. ‡automotive cockpit POV environment.

		Symbol detection accuracy: mean (μ) and variance(σ^2), n = 2						
		control series		**low-load series**		**high-load series**		
trial	**symbols/trial**	μ	σ^2	μ	σ^2	**symbols/trial**	μ	σ^2
1	24	0.688	1.05E-01	0.979	8.41E-04	36†	0.694	2.46E-02
2	24	0.938	7.74E-03	0.979	8.41E-04	45†	0.923	2.89E-04
3	24	0.958	3.48E-03	0.938	7.74E-03	45†	0.956	1.00E-06
4	24	0.958	3.48E-03	1	0	18‡	0.972	1.52E-03
5	24	0.958	3.48E-03	0.979	8.41E-04	18‡	0.972	1.52E-03
6	24	0.917	1.39E-02	1	0			
7	24	0.938	7.74E-03	0.979	8.41E-04			
8	24	1	0	1	0			
9	24	1	0	0.979	8.41E-04			
10	24	0.979	8.41E-04	1	0			

6.5 Detection speeds and learning trends

6.5.1 Control Environment

Longitudinal representations of each subject's performance in the control series facilitate an assessment of how detection speeds evolve over repeated exposure to the codex. Detection speed performance, as well as, raw detection timed responses were filtered for absent or incorrect symbol identification. As summaries of symbol detection accuracy from the previous section demonstrate, underlining excluding these points from the data set did not greatly reduce the total number of data points. For each subject, a nonlinear model of an exponential decay was fit to detection times. Raw data with fitted curves and their coefficients are shown in Figure 31 spanning the following two pages.
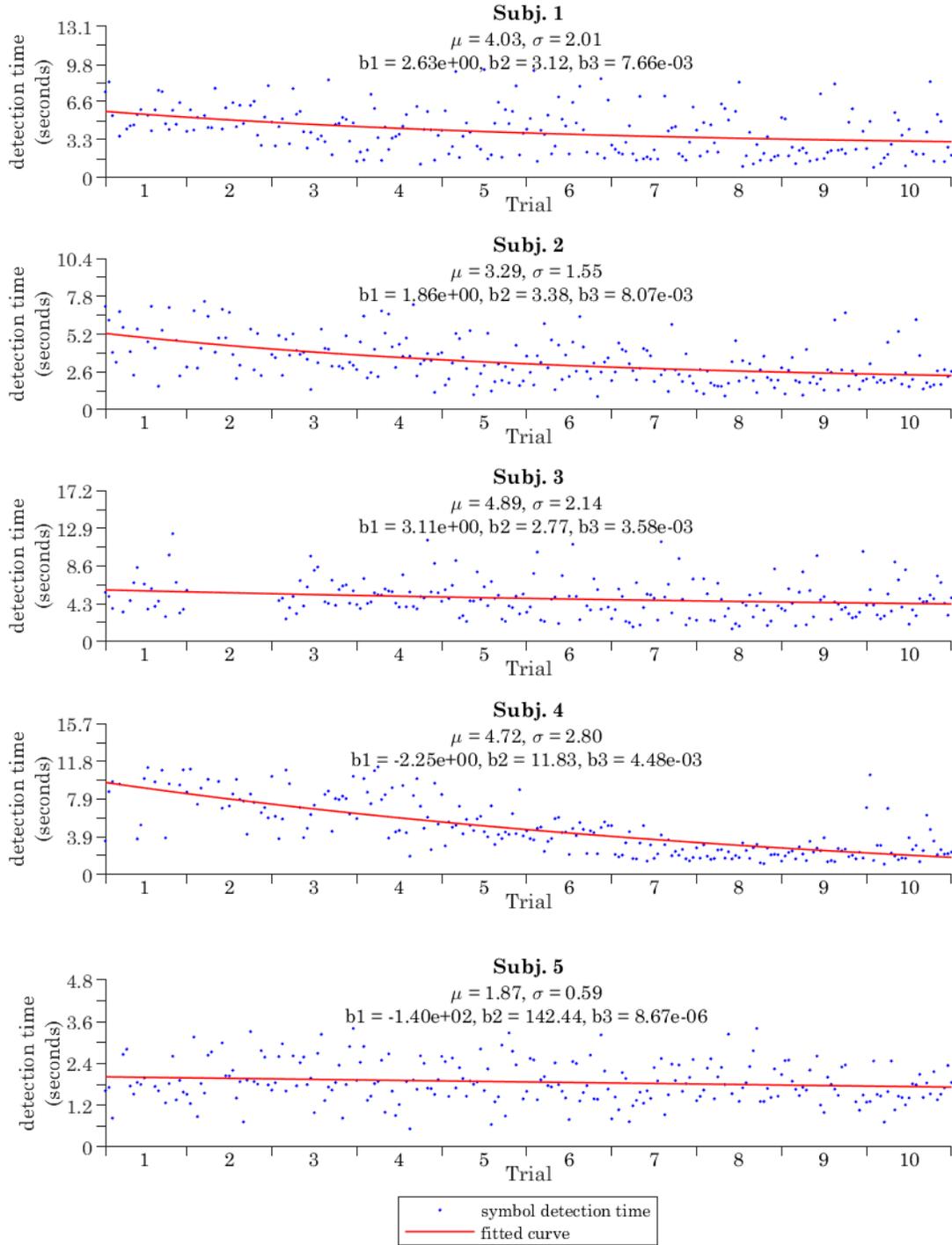
Figure 31. (A) Symbol detection rates for subjects {1,2,3,4,5} displayed longitudinally over 10 successive trials in the control series (standard psychophysical environment). Coefficients of the exponential decay model $\{b_1, b_2, b_3\}$ are indicated for each subject, as well as mean (μ) and standard deviation (σ).
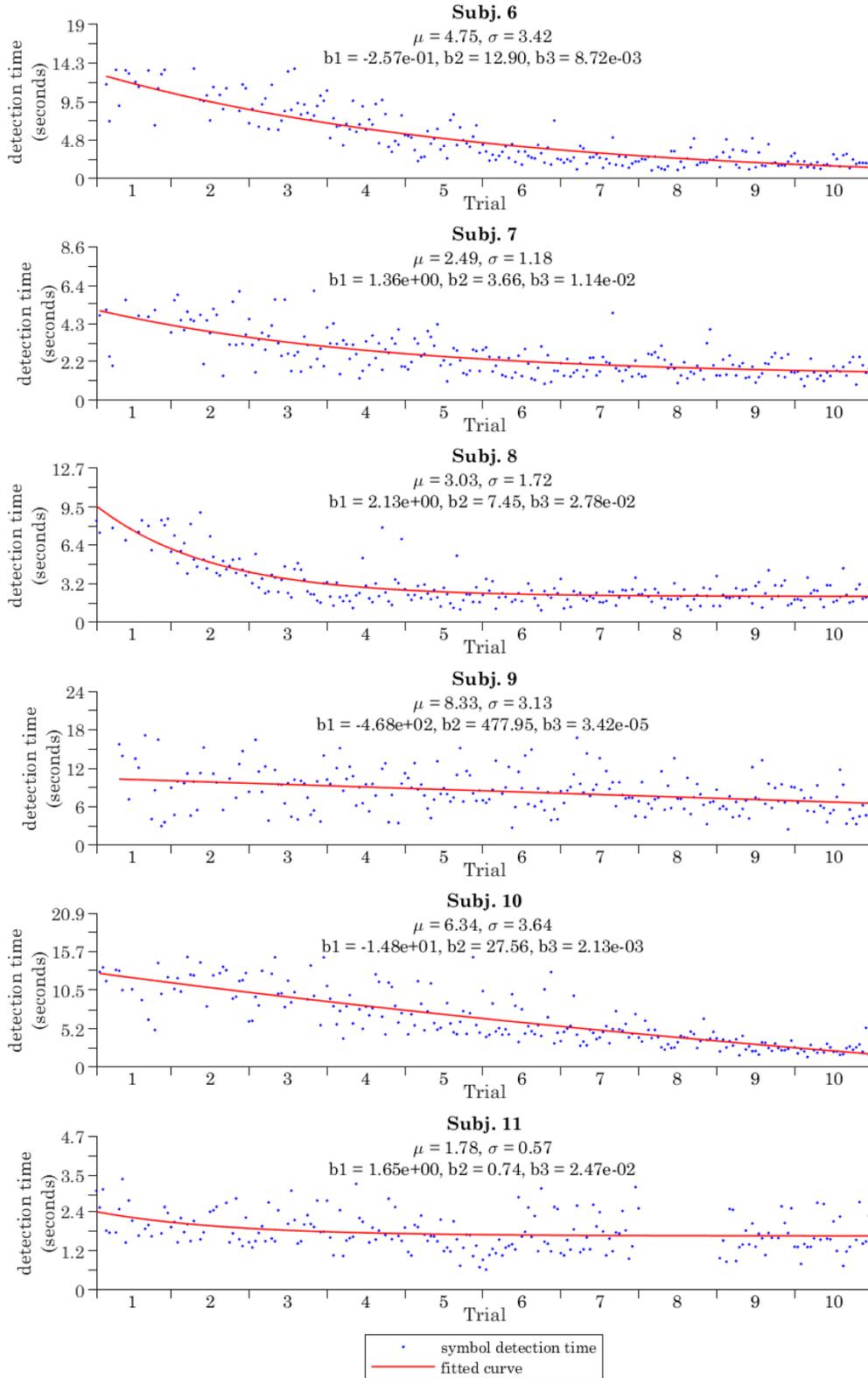
Figure 31. (B) Symbol detection rates for subjects {6,7,8,9,10,11} displayed longitudinally over 10 successive trials in the control series (standard psychophysical environment). Coefficients of the exponential decay model $\{b_1, b_2, b_3\}$ are indicated for each subject, as well as mean (μ) and standard deviation (σ).

6 | Results

6.5.2   Longitudinal: all series

A subset of subjects completed multiple series: five subjects (Subj. 1,2,3,4,5) completed high-load trials from <u>automotive cockpit POV</u>. Before completing high-load trials, two subjects within this subset (Subj. 1,2) completed <u>low-load series and three additional high-load trials from pedestrian POV</u>. Results from Subj. 1 and 2 who completed all series (control, low-load, and high-load) are represented longitudinally below in Table 8 and Figure 32.

As in the control series longitudinal analysis, trends in visual learning were characterized by fitted curves. A nonlinear model of exponential decay was fitted to detection times in control and low-load series while a linear model was fit to high-load data.

Control series data from Subject 1 below exhibits some non-linear characteristics but low-load data is functionally linear, albeit still trending downward. High-load series trends exhibit little change over time, as evidenced in small p_1 values for both pedestrian and automotive cockpit POV trials. Using linear trends in high-load cases for any predictive assessment should be done cautiously due to both greater variance in data and smaller sample size. Total sample sizes for the three series are 240 symbol detection times in the control, 240 in low-load, 126 in the high-load pedestrian POV, and 36 in the high-load automotive cockpit POV.

Table 8. **Subject 1 longitudinal series metrics**: mean (μ) and variance (σ^2) of detection times across all symbols within each study (with equations and coefficients of fitted curves)

Series	μ	σ^2	curve eq'n	coefficients	RMSE
Control	4.03	4.040	$y(x) = b_1 + b_2 * e^{-b_3 * x}$	$b_1 = 2.63, b_2 = 3.12, b_3 = 7.66\text{e-}3$	1.88
Low-load (pedestrian)	1.82	0.980	$y(x) = b_1 + b_2 * e^{-b_3 * x}$	$b_1 = -487.76, b_2 = 490.53, b_3 = 0$	0.838
High-load (pedestrian)	2.83	1.588	$y(x) = p_1 * x + p_2$	$p_1 = -1.63\text{e-}3, p_2 = 2.92$	1.27
High-load (driving)	3.24	1.877	$y(x) = p_1 * x + p_2$	$p_1 = -6.38\text{e-}3, p_2 = 3.35$	1.39
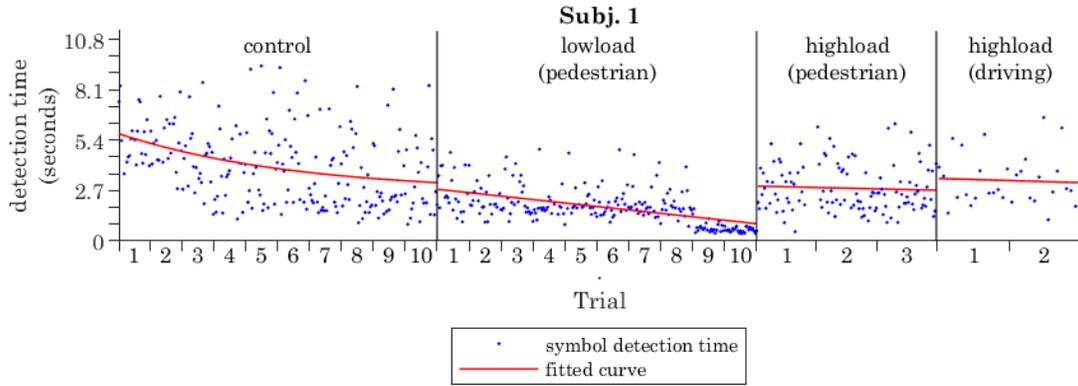
Figure 32. Symbol detection speed over successive trials in control, low-load, and high-load series for Subject 1. Trends in control series are fitted to an exponential decay trajectory while subsequent series (low-load and high-load) exhibit linear characteristics.

Longitudinal analysis of Subject 2 throughout control, low-load, and high-load series is represented below in Table 9 and Figure 33. Similar to previous analysis for Subject 1, Subject 2 exhibited exponential decay trends throughout the control series and functionally linear trends in detection speeds in the low-load case were recorded. Variance in the high-load case increased and linear trends had similarly small p_1 values.

Table 9. **Subject 2 longitudinal series metrics:** mean (μ) and variance (σ^2) of detection times across all symbols within each study (with equations and coefficients of fitted curves)

Series	μ	σ^2	curve eq'n	coefficients	RMSE
Control	3.29	2.403	$y(x) = b_1 + b_2 * e^{-b_3 * x}$	$b_1 = 1.86, b_2 = 3.38, b_3 = 8.07\text{e-}3$	1.34
Low-load (pedestrian)	1.71	0.292	$y(x) = b_1 + b_2 * e^{-b_3 * x}$	$b_1 = 1.38, b_2 = 0.762, b_3 = 7.78\text{e-}3$	0.512
High-load (pedestrian)	2.56	1.277	$y(x) = p_1 * x + p_2$	$p_1 = 1.53\text{e-}3, p_2 = 2.45$	1.13
High-load (driving)	1.64	0.348	$y(x) = p_1 * x + p_2$	$p_1 = \text{-}1.09\text{e-}3, p_2 = 1.66$	0.600
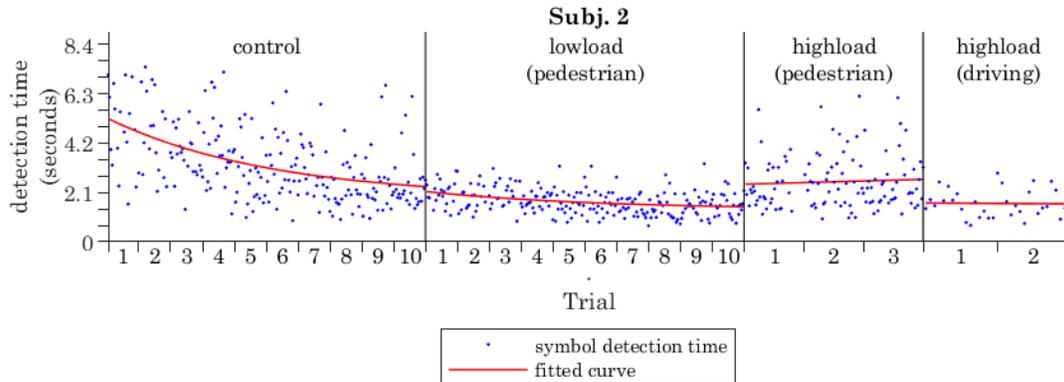
Figure 33. Symbol detection speed over successive trials in control, low-load, and high-load series for Subject 2. Trends in the control series are fitted to an exponential decay trajectory while subsequent series (low-load and high-load) exhibit linear characteristics.

6 | Results

6.5.3 Longitudinal: control to high-load automotive cockpit POV

A key question the studies aimed to address was how well training across the control series prepared observers to quickly and accurately detect motion-modulated symbols in the far periphery. A second first batch of three subjects completed the control series and high-load automotive cockpit POV trials, without intermediary training on low-load series and high-load pedestrian POV environment. Longitudinal results are represented in Figure 34 below. Trial mean, variance, fit coefficients, and RMSE are summarized in Table 10.

.
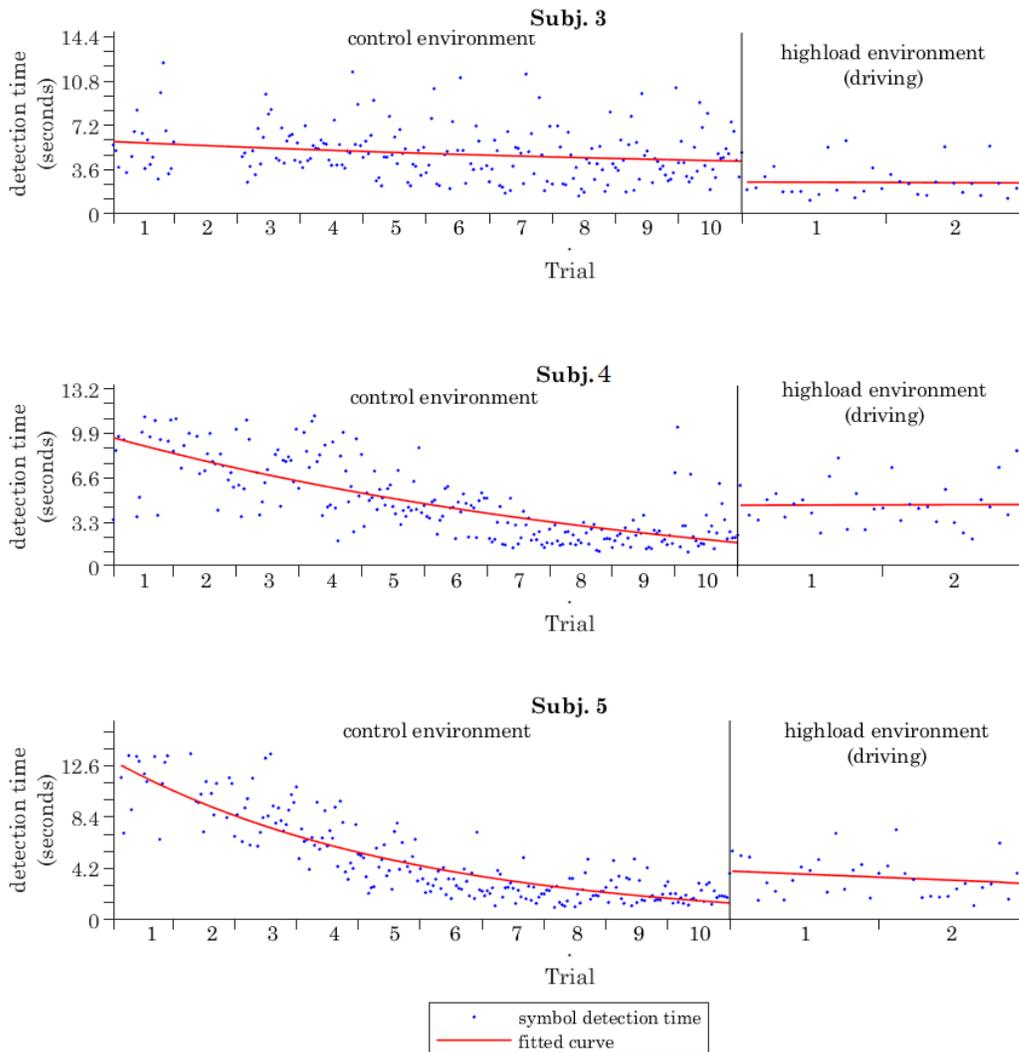
Figure 34. Symbol detection speed over successive trials from control and high-load driving series for subjects {3,4,5}.

Table 10. Longitudinal metrics for control and high-load series (subjects 3,4,5): mean (μ) and variance (σ^2) of detection times across all symbols within each series (with forms and coefficients of fitted curves)

Subject	Control series					High-load driving series				
	μ	σ^2	fit	coeff.	RMSE	μ	σ^2	fit	coeff.	RMSE
3	4.89	4.580	Exp. decay	$b_1= 3.11$, $b_2= 2.77$, $b_3= 3.58\text{E-}3$		5.332		Linear	$p_1=$ $p_2=$	
4	4.72	7.840	Exp. decay	$b_1=-2.25$, $b_2=11.83$, $b_3= 4.48\text{E-}3$		3.080		Linear	$p_1=$ $p_2=$	
5	4.75	11.696	Exp. decay	$b_1=-0.257$, $b_2=12.90$, $b_3= 8.72\text{E-}3$		3.838		Linear	$p_1=$ $p_2=$	

To compare improvement outcomes from control series to high-load driving trials, a normalized improvement factor was calculated. First, control series trail 1 detection time mean was determined for each subject, serving as a benchmark for un-trained symbol detection performance. Next, detection time mean across both trials of the high-load driving series was calculated and normalized to control series trial 1 mean to arrive at the normalized improvement factor. These values are summarized below.

Table 11. Mean detection time performance in control series trial 1 and high-load driving trials 1-2, yielding normalized improvement factor f_i indicating relative changes in detection time.

Subject	Control series trial 1 mean μ_{C1}	High-load driving series trials 1-2 mean μ_{HL}	normalized improvement factor $f_i = (\mu_{C1} - \mu_{HL})/\mu_{C1}$
1	6.497	1.897	0.708
2	6.446	3.791	0.412
3	8.889	5.332	0.400
4	10.764	3.080	0.714
5	15.146	3.838	0.747

The two groups of high-load driving participants who either completed increasingly complex trials prior to the high-load case (subjects 1 and 2) or completed no additional training (subjects 3, 4, and 5) exhibited no significant difference in normalized improvement factors (p = 0.76).

7. Discussion

This dissertation proposed a new methodology for computational generation and delivery of complex information to far eccentricities of the human visual field. Systematic and applied longitudinal studies across a population of subjects yielded significant preliminary findings demonstrating:

- accurate complex symbol discrimination over short (< 2 seconds) timescales
- trends in accurate symbol detection speed over successive trials support early learning adoption of a new visual codex
- far peripheral complex symbol recognition in highly dynamic visual environments absent of central fixation, fixed contrast, luminance and illuminance
- Symbol detection rates visual learning outcomes persist with increasing complexity in the visual environment
- framework and methods for delivering semantic information to far peripheral regions of the human retina are a valuable extension of contemporary methodologies

Results from low-load and high-load studies with increasingly complex environments (absent of central fixation and fixed contrast/luminance/illuminance) show comparable rates of complex symbol discrimination. Most notably, improvements in detection rates achieved through visual learning in the standard psychophysical environment are also present in low- and high-load studies at varying offsets.

7.1 Symbol detection rates correlate with motion energy with notable exceptions

Motion energy analysis provided an objective, quantitative metric by which each of the codex blocks could be evaluated (see Figure 35 below). This representation shows performance of each symbol in detection speed across the control environment versus absolute motion energy. A linear trend is fit to approximate the relationship between absolute motion energy E_{total} and steady-state detection rate (calculated from the final trial in the control series, n=11). Absolute motion energies of higher complexity symbols such as *stick figure* and *boat* were greater than simpler geometries such as *4* or *C* which is consistent with the expectation that symbols of higher-order (more apertures) and longer duration convey more

motion energy.   What's interesting in this representation is a noticeable cluster of recognition times (in seconds) across the majority of symbols presented in the control series. In terms of spatial differentiation of time over motion energy, two significant outliers deviate from the approximate trend. A representation of absolute motion energy versus detection time is shown below, with symbols demarcated by their order (1st through 4th order, indicating the number of apertures present in the codex block).
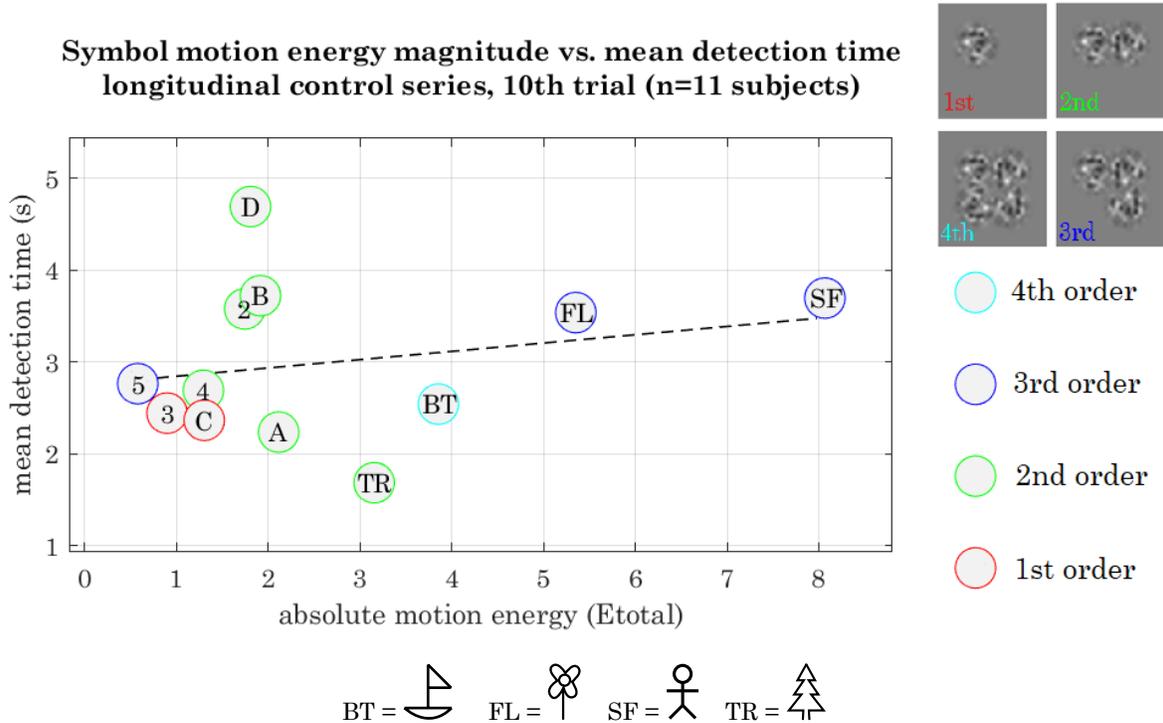
Linear model	p_1	p_2	RMSE
$y(x) = p_1 * x + p_2$	0.0904	2.755	0.8637

Figure 35. Spatial differentiation of detection rate over motion energy, with symbol order indicated by marker edge color (red – 1st order, green – 2nd order, blue – 3rd order, teal – 4th order). The linear trend is indicated by the dashed line. Mean detection time in seconds is calculated from the final trial of control series for all subjects (n=11), representing the closest approximation of steady-state detection rate after visual learning has occurred.

Notable outliers above, *D* and *TR* (shorthand for *tree*), while both 2nd order symbols, fall on either side of the trend. Improvement in detection rates observed for *TR* is best explained by its' uniqueness in the codex – *TR* was the only symbol to present <u>motion stimulus in multiple apertures simultaneously</u>. Improvements in detection times over repeated trials could be attributed to many factors. One of the most significant is the ability

to discriminate symbols in the context of a known subset. In this respect, the uniqueness of *TR* contributed significantly to its low mean detection times.

Uniqueness also played a critical role in the outcomes for symbols *D* and *B*. While having comparatively low motion energies, these required the most time for accurate detection. In terms of stroke trajectories, *D* and *B* are <u>functionally identical for the first half of symbol presentation</u>. This lack of uniqueness imposes a limit on improvements in detection rates for these symbols. Furthermore, we know from prior work in parameterizing MIPS stimuli, *tangential motion*, such as the vertical downstroke of *B* and *D* symbols, is far less pronounced than horizontal strokes moving toward or away from the fovea [26], and could have also impacted the perceptibility of these codex blocks.

Uniqueness as a determinant of detection rate does not account for the discrepancy between *B* and *D*, however; to account for this, we must consider the impact of features within white Gaussian noise grain. When grain moves behind an aperture along the stroke path, patterns and features in the grain can create small apparent motion effects. In the case of *D*, an anomaly in the grain causes the downward stroke to appear wavy. These small motion effects were sufficient to cause confusion in the first half-second of symbol perception, greatly limiting improvements in detection time for *D*. For future codex development, one possible approach to identifying unwanted motion effects from grain anomalies could include comparing the desired motion with optical flow measurements.

The role of confusion (i.e. false positives), while not formally evaluated in these experiments, is a critical consideration in the design of a visual language. Within the results of the experimentation described here, there were several recorded instances of false positive symbol detection between the characters *B* and *D*, but not between any other codex blocks. While there are other instances of similar stroke geometries, such as *3* and *2*, these blocks were distinct in other easily perceptible ways (i.e. number of apertures or codex block duration). Only symbols *B* and *D* were similar across stroke geometry, number of apertures, and codex block duration. In future studies to refine and improve the codex, a key component of optimization will be the design of a library to maximize uniqueness across these three parameters.

Ultimately, the majority of symbols fall within a time frame for complex semantic information delivery that may have significant impact in industry, both in aerospace and

aeronautical instrumentation design, but illustrate a positive trend in the development of generating new stroke geometries to further refine a broader development of codices.

7.2 High accuracy rates are achieved quickly and maintained across environments

One of the most exciting outcomes from data captured across all subjects and symbols was the high rate of detection accuracy. Notably, all subjects reported with 100% accuracy in the 9th trial of the control series. Furthermore, 95% of subjects reported with at least 87.3% accuracy by the second trial, which is very encouraging for the applicability of this approach with minimal training.

Beyond initial demonstration of detection in controlled environments, studies aimed to assess the feasibility of far peripheral semantic information delivery in increasingly complex visual scenarios. Low-load series expanded upon the control environment by presenting a dynamic pedestrian POV. The same video was used for each symbol, to provide a baseline independent of symbol characteristics for detection accuracy in dynamic environments. Results show a negligible reduction in detection accuracy when transitioning from a 50-gray background with static fixation cue (control environment) to a natural scene with uncontrolled changes in brightness, hue, and contrast. High-load series expanded this approach to evaluate observer performance when symbols are presented in random sequence, without pausing for reporting, throughout a continuous video lasting several minutes (video durations from different trials ranged from 3:29 to 8:08). In the first high-load trial, mean detection accuracy rates dropped to levels comparable with the first trial in the control series, but rebound just as quickly. High (>90%) accuracy rates are maintained after the first high-load trial throughout the series, increasing steadily despite significant increases in environment complexity from trial 3 (pedestrian POV environment, town center) to trial 4 (automotive cockpit). This suggests the new trial format of symbols presented over a continuous video source (with no pause for reporting), contributed more to lower accuracy rates in the first high-load trial than the complexity of the visual environment itself.

7.3 Near-minimal detection speeds are achievable for the majority of subjects

Detection time required to make an accurate identification of a complex symbol is a critical metric in assessing practicality and feasibility. Observers who participated in the control series completed 10 trials (each trial with a re-randomized sequence of symbols but otherwise identical) to determine potential for improving detection speeds through visual learning. Figure 31 on pages 61 and 66 shows detection times for each subject arranged longitudinally over ten successive trials with fitted curves to demonstrate trends in detection speeds. These fitted curves are assembled in the figure below (Figure 36) with the accompanying table to summarize these trends.
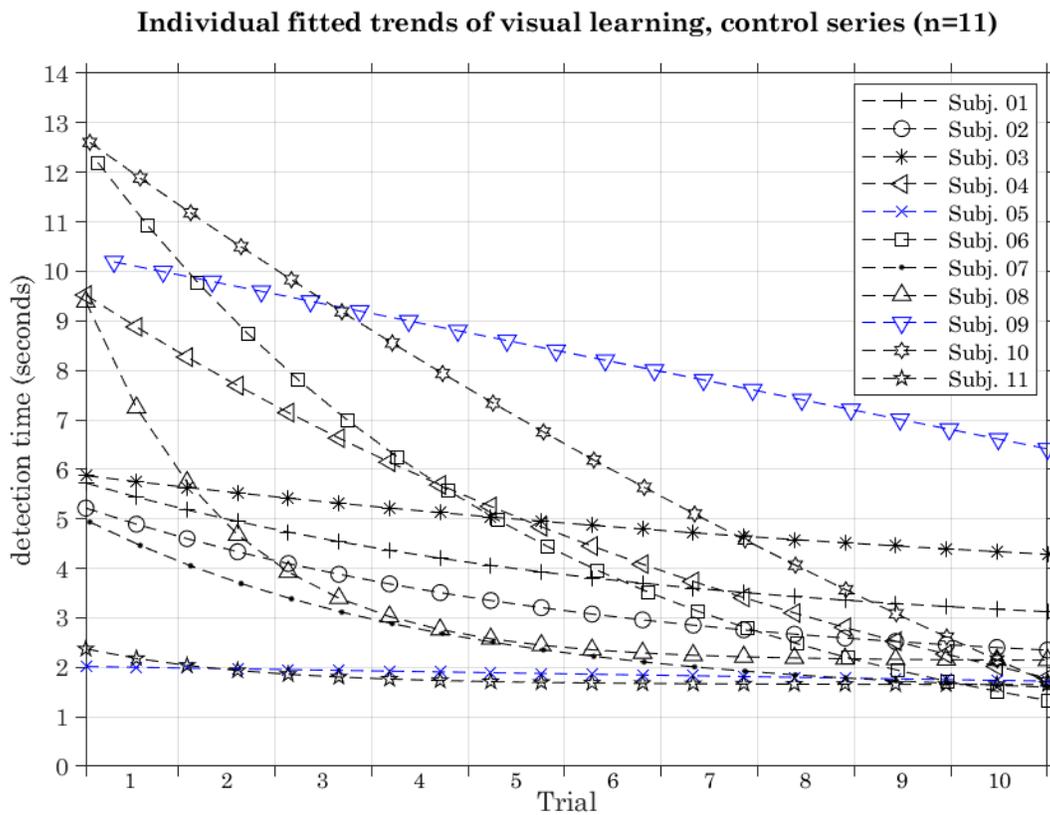
Figure 36. Fitted trends representing improvements in detection speeds over successive trials for each subject in the control series (n=11). An exponential decay model was attempted (black); if the third coefficient p_3 dropped below a certain threshold (0.001) a linear model was fit to data (blue).

Table 12. Fit data summary for control series (n = 11). Exponential decay (exp. decay) trends were fitted to the non-linear model $y(x) = p_1 + p_2 e^{p_3 x}$; linear models were fitted to $y(x) = p_1 x + p_2$. Root mean square error was normalized to the span of the data for each subject (NRMSE). Steady-state approximations represent the rightmost and minimum value of each fitted trend – these values are more predictive for some subjects than others depending on the model's rate of decay (in the nonlinear case) or slope (in the linear case).

subject	model	p_1	p_2	p_3	NRMSE	accurate symbol detection rate steady-state approximation (s)
1	exp. decay	2.63	3.12	7.66E-03	22.2%	3.12
2	exp. decay	1.86	3.38	8.07E-03	20.5%	2.35
3	exp. decay	3.11	2.77	3.58E-03	16.5%	4.28
4	exp. decay	-2.25	11.83	4.48E-03	18.4%	1.79
5	linear	0.00	2.02	0	20.3%	1.72
6	exp. decay	-0.26	12.90	8.72E-03	11.4%	1.33
7	exp. decay	1.36	3.66	1.14E-02	16.1%	1.60
8	exp. decay	2.13	7.45	2.78E-02	12.6%	2.14
9	linear	-0.02	10.32	0	20.1%	6.41
10	exp. decay	-14.84	27.56	2.13E-03	15.2%	1.71
11	exp. decay	1.65	0.74	2.47E-02	19.8%	1.65

Making comparisons between subjects should be done cautiously, as studies did not control for all possible variations between observers beyond consistent environmental conditions and visual acuity. That being said, Figure 36 demonstrates two significant outcomes from results.

First, improvements in steady-state detection speeds across subjects appear to reach a floor around 1500 ms. This is likely a feature of design characteristics of the codex blocks. Symbol stroke speeds determine how long it takes observers to confidently discriminate one symbol from another and specific stroke timing of this codex limits how quickly similar symbols (such as *3* and *2*) can be differentiated. True lower limits of accurate detection speeds could be evaluated in future studies by increasing symbol stroke speeds until a falloff in detection accuracy occurs.[10]

---

[10] There is likely a theoretical limit to detection speed on the order of 200-500ms which reflects the timescales of the many complex processes which manifest as visual symbol recognition [44].

7 | Discussion

Second, near-minimal detection speeds were achievable for subjects with both fast and slow initial speeds. Minimum values in trends were between 1300 and 2350 ms for eight out of eleven subjects, including those with slowest initial speeds. Overall, exponential decay models demonstrate that repeated exposure to the codex greatly improves detection speeds in all subjects in the control series.

7.4 Longitudinal outcomes in dynamic environments

A subset of participants from the control series continued to complete trials in increasingly complex visual environments. An additional (n=2) completed a low-load series and two high-load series (from pedestrian continuing through automotive cockpit POV trial sequences), while another group of participants (n=3) completed the automotive cockpit POV, in addition to the control. Comparing these two sub-subject groups allows us to evaluate whether continued exposure to the codex in environments, with gradually increasing complexity, resulted in improved detection rates.

7.4.1 Detection speeds continue to improve despite increasing scene complexity

Both subjects who completed the low-load series in the pedestrian POV environment continued to improve in reported detection times over successive trials, despite the dramatic shift from a static 50-gray background and static fixation to a dynamic natural scene. Notably, fitted trends (in Figure 32 and Figure 33 for Subjects 1 and 2 respectively) show close continuity throughout the series, especially near the end of control trials and beginning of low-load series.

In addition to continued downward trends in the low-load case, both subjects exhibited a substantial decrease in variance of recorded detection times. These outcomes from the low-load series represent the first demonstration of far peripheral complex symbol recognition in dynamic visual environments. High accuracy rates and continued improvements in detection speed in low-load series build a strong foundation for feasibility and practicality of implementing motion-modulated peripheral stimuli for semantic information delivery outside of controlled psychophysical environments.

High-load series results for this group show an increased variance in detection times when compared to low-load series, which is to be expected when transitioning from a repeated

dynamic environment with pausing for reporting to a continuous stream of presented symbols throughout a dynamically changing natural scene. However, detection speeds in both pedestrian and automotive cockpit POV trials still benefitted from prior training, with 71% and 41% improvement in mean detection times from the first control trial to high-load driving trials for subjects 1 and 2 respectively.

7.4.2 Visual learning outcomes from control series persist in complex environments both with and without additional exposure

To evaluate the extent of visual learning accomplished by the control series, three additional subjects completed high-load driving trials, *without* any additional studies in between. In Figure 34, an exponential curve was fit across all three subjects for the control series (as it was for the rest of the control subjects). In addition, you can see the linear fitting of all three subjects having completed the high-load driving environment. What is interesting to note is DC offset across all three subjects in the high-load is significantly lower than initial detection rates in the control series. While three subjects are not enough to quantifiably determine the meaning of their offset detection times between control and high load, it is however sufficient to make a supposition toward design of further studies to refine when optimal learning is achieved. While it is outside the scope of this dissertation to draw conclusions comparing these three subjects (who did not complete low and high-load pedestrian settings) it is fascinating to draw an inference between the leveling off of average mean detection times weighed against the other two subjects (see page 64). Further studies are necessary to quantify DC offsets in varying conditions after learning has been achieved.

Improvement factors calculated for each subject show no significant difference in improvements between subjects who completed the entire study series and those that only completed control and high-load trials (see Table 11). Similar performance between these groups suggests sufficient priming on the codex occurred during the control series to prepare subjects for dynamic visual environments beyond the standard psychophysical backdrop. These initial findings illustrate a significant potential toward establishing a visual learning model with a broader group of subjects in subsequent trials, yielding a more complete characterization of training necessary for achieving high symbol detection accuracies in real-world scenarios.

7.5 Limitations

Human capacity for processing, parsing, and evaluating meaning in dynamic visual environments is subject to the myriad signal potentials which it receives in a constant stream of sensory perception. How acute is the bandwidth of our selective attention within the confine of a singular sensory mechanism? How well can we process information coming from one sensory modality and ignore the perceptual dimension of others?

In any experimentation, there are uncontrollable sources of variation – the noise floor set by practical realities of study design. When evaluating the manifestation of a perceptual phenomenon, studies must account for the qualities of human sensory processing which are a constant: adaptation, multisensory integration, and subject uniqueness.

Psychophysical experimentation must always balance the importance of repeatability with the inevitability of adaptation. A single trial in the control series consisted of 24 codex blocks so that throughout an entire series of 10 trials a subject completed 240 symbol detection tasks in sequence. Random order of codex block presentation is sufficient to prevent adaptation to the unique stroke geometries of each symbol; however, it is highly likely that the rhythm of repeated trials contributes in some part to improvements in detection speed and accuracy. Extending the term of longitudinal trials, to encompass a longer duration of time between sessions, could more clearly elucidate the role of visual learning in improved detection rates over time.

Quantizing visual perception as an expression of experience is referentially viewed as a modular function, temporally sequenced within multiple sensory inputs that operate largely as separate and independent channels in the first cascade of interaction. Design characteristics of the simulated environment isolated and targeted visual processing as much as possible, removing opportunities for cross-talk. Throughout administration of the trials, any codex block presentations in which distractors were present (i.e. audio pollution or technological interruptions) were marked and excluded from the analysis.

Finally, an individual's sensory sensitivity is inherently biased by any number of environmental factors – caffeine levels, sleep patterns, stress levels, etc. By maintaining a consistent environment across participants, we assume contribution to the noise floor from subject uniqueness is approximately Gaussian.

Gaze tracking required subjects to wear a headset, while the application of the potential of this technology may be limited in real-world environments for individuals who wear glasses. Due to the nature of the experimentation, the headset was necessary for calibration and verifying the results: gaze during stimulus presentation was collected and evaluated for any deviation outside 30 degrees from the presented central fixation point. However, fixations within the 30-degree arc were not evaluated and no differentiation was made between monocular and stereoscopic vision. While outside of the scope of establishing baselines for this dissertation, it is a critical component to weigh against perceptual outcomes and development of user interfaces.

7.6 Future work

Findings presented here illustrate high rates of detection and accurate recognition of complex symbols beyond 50° eccentricity throughout longitudinal control, low-load, and high-load series across vector geometries of generated characters. Completed experimentation establishes a foundation for a new paradigm of peripheral information delivery and, while purposefully limited to first principles, it carries potentials to exceed the scope of single codex presentations. Immediate relevant avenues for exploration include cognitive load assessments via applied psychology protocols and contextualizing visual perceptual outcomes demonstrated here within other sensory mechanisms, particularly aural modalities. Gaze data, implemented for filtering out responses in which gaze was diverted, could be integrated into a formal analysis of trends in fixation throughout stimulus presentation. For example, future studies could establish causal links between stimulus onset and gaze trends. Such analyses would provide an important dimension to evaluating the impact of stimulus delivery on fixation during high-load tasks. Beyond explorations of attention and gaze, notable opportunities for future work are summarized below.

7.6.1   Further parameterization of stroke characteristics and perceptual outcomes

Outliers *D* and *TREE* are a clear example mean across subjects for how calibration and a complete library of rotational vectors can be generalized, highlighting the need for further study in defining and optimizing a programmable manual for individual stroke paths. What are the most efficient perceptual stoke lengths weighed against high accuracy responses? How do perimeters of persistence of vision weighed against grain and envelope manifest in

accuracy and speed of detection? As noted in the discussion, measuring optical flow of each codex block could better quantize mechanisms of motion that yield perceived position shifts. Further work is needed to calibrate stroke path curvature and aperture placement to perceptual outcomes. Fully parameterizing symbol motion geometries and their manifestations could facilitate the development of an extensive library of stroke vectors weighed against grain size, resolution, and aperture placement (both spatially and sequentially in time).

Further, I can only theorize that, given the outcomes of this work, a significant series of studies need to be performed to weigh reception of individual stroke trajectories as they relate in combination in generating complex semantic delivery, i.e. why in second-order (two aperture) codex blocks when the first stroke is vertical, as in symbol *D*, learning adoption rate is low relative to time? How does this relate to fourth-order carriers, such as symbol *flower*, with greater absolute motion energy (number of strokes) and higher adoption rate for learning? While outside of the scope of this initial work, there is a significant and exciting opportunity to categorize and develop a computational method for extending a library of motion-stoke vector geometries for future applications in delivering complex semantic information.

7.6.2   Comparing horizontal presentation of far peripheral codex cues.

Over the course and duration of longitudinal studies, from control to low-load to high-load in pedestrian and automotive cockpit POV scenarios, cues were horizontally presented in the same visual axis across symbol presentation and trial. This was necessary to establish a framework for repeatability to be able to quantize findings across individual performance. Moving forward, studies need to be designed and conducted to evaluate detection speed and accuracy rates of presented peripheral stimuli across vertical zonal regions.

Development and initial placement implementation of the codex cues were informed by early experimentation and observational findings which support the theories of the evolution of receptive zone distribution and sensitivity in human vision. Practically, this manifests in greater sensitivity to motion modulation in the upper region of our far periphery than the lower [75]. We are more sensitive to objects falling toward us from the sky than to the same motion at our feet. That is not to say highly efficient peripheral motion envelopes will not be

as successful in lower visual regions. It is a fascinating further avenue to explore and opens perhaps one of the most significant future research and outcomes of this work.

7.6.3   Quilted aperture arrays

High accuracy adoption rates in both the time and physio/spatial domains across subjects were demonstrated through findings and portray an exciting opportunity to further multiplex information delivery through established techniques. If we can recognize a complex symbol, such as *boat* or *stick figure* within a persistence of vision stroke threshold (500 ms/stroke path), can multiple complex symbols be presented simultaneously in an array of static apertures to far peripheral regions of the visual field?

Creating further spatial envelopes beyond fourth-order is perhaps the single most important and personally fascinating research avenues for continued exploration.   Protocols implemented to establish first principles were limited in the scope of experimentation, i.e. single linear symbol presentation.   However, it is my firm belief that we need to lay a computational framework and delivery mechanisms for what I would call **quilted aperture arrays**: clusters of many apertures (> 20) synchronously or sequentially mapping perceived position shifts to a manifestation of textures, multidimensional surfaces, or high detail imagery.

7.7 Applications

Societal and ethical considerations for future implementation are at the forefront of potential applications for systems described in this dissertation. Eyesight is a rapidly and dynamically changing sensory mechanism that we possess, and throughout our lives its power and ability naturally increases and declines. In those medians, we are charged with the responsibility of leveraging all of the knowledge and technological advances we possess to make a better future. Our safety, our health, and our interconnectivity benefit from building platforms for early learning with disability, early alert potentials, safety, navigation, and new avenues in augmented reality and media formats.   The systems and results presented in this dissertation serve as a foundation, establishing first principles of this approach. Next steps outlined in the previous section on future work are critical to its continued development, toward implementation in technologies and environments with the greatest potential for

multiplexing visual information delivery. Notable opportunities for deployment of far peripheral semantic information delivery are detailed below.

State of the art in early (real-time) visual reporting in high-risk environments is delivered through multiple metrics to an operator: physical instrumentation overlaid with heads-up displays (HUDs) and attended with audio cues and other situational informatics experienced and programmed through multi-sensory feedback. I believe there is a significant potential benefit supported by findings defended in this dissertation lending toward redistributing cognitive and visual load to non-task dividing technologies to improve performance awareness in highly dynamic, demanding environments.

Instrument control engineering for aerospace technologies, such as human flight assistance, control deck, pose some of the most exciting opportunities for implementation of methodologies presented, especially in low gravity and high-G environments in which intraocular pressure is not a constant. Documented reductions in near-eye visual acuity during long-term space flight [76] are a critical implication for the usability of near-eye displays in microgravity. The approach described in this dissertation is not reliant on high visual acuity for the communication of complex symbols and by leveraging peripheral vision alone for this task it is highly robust to such changes in vision. To this end, further exploration in integrating existing real-time adaptive projection systems (HUD and far peripheral display platforms such as those capable of delivering flicker) would prove to be invaluable in accelerating the evaluation of immediate application in both consumer and industrial domains.

Head-mounted displays, whether whole-field or peripheral, are constrained by limits of hardware and pixel real estate available on a mobile platform. Yet, the scope of applications enabled by these technologies spans an incredibly exciting space of high cognitive load environments. As advances in hardware continue to expand accessible field of view, these technologies are a promising use-case for the deployment of the systems and methodologies described in this dissertation.

Moving beyond digitally integrated envelopes embedded onto display technologies (such as monitors, or headset), there are significant opportunities to leverage standalone devices placed in natural environments. Most driver cues for road safety are presented peripherally to the operator. As many traffic lights have been replaced with LED arrays, and stop signs

illuminated with such, integrating Gabor structured motion into their function could dramatically increase driver awareness beyond static illumination, retroreflection, or flicker. At low computational and hardware costs, such interventions could contribute to a societal shift in the way we develop and confront commercial and domestic transportation. Perhaps the most impactful feature afforded by the systems described here is the extent of perceived shift achievable within a small, static aperture. As mentioned previously, an important avenue of future work is evaluating the role far peripheral symbol recognition plays in attention, and a further facet of future work should distinguish this approach from flicker and other established notification methods. With further parameterization, the expanded perceptual outcomes achievable with perceived position shifts could be leveraged in many further applications, including:

- pedestrian crosswalk indicators
- bike lane demarcation
- running lights and breaking alerts
- air traffic control and programmable assets
- tactical communication in nonverbal high load environments (aerial and subsurface)
- worker safety in commercial settings
- environmental augmentation (urban and home environments)

One of the greatest motivators behind my research, both personally and academically, is to leverage engineering for the health, safety, and betterment of society. Activating peripheral vision for complex information delivery has exciting implications in the scope of:

- learning models for education, ADHD, and cognitive multiplexing
- applications for uncorrected vision, low light environments
- learning with visual disabilities: reading from peripheral vision both for early-onset childhood cataracts and degenerative vision due to glaucoma, macular degeneration, cataracts, and clinical applications

Future work in the area of adaptive stimulus generation will expand to cross-modal sensory experience/integrations where it may have greatest relative value as an experiential output to the observer, in such areas as neural mismatch models, in the affects, control, and mitigation of kinetosis. Beyond kinetosis, there are immediate potentials for deploying

motion modulated proprioceptive and semantic cues into dynamic environments in a myriad of applications that I sincerely hope this work inspires.

8. Conclusion

The goal of this research was to challenge established limits of complex symbol recognition by developing a novel codex leveraging the unique faculties of peripheral visual processing. Instead of relegating peripheral vision to the role of intermediary between attention and context, the systems described here aimed to activate unattended channels of information delivery with a new visual language, encoded in the functional biology of peripheral mechanisms.

This approach expands upon prior work not only in first principles of information composition and delivery but also in its capacity to convey highly abstract forms via apertures spanning only 0.64 degrees of the visual field. This spatial compression of complex semantic information has immediate applications in constrained display environments such as virtual and augmented reality.

By integrating powerful perceptual phenomenon against dynamic and real-world environmental constraints, this investigation proposed, as a foundational framework, a novel systems architecture for generation and adaptation of stimuli, peripherally presented to discrete regions of the human visual field. Methodologies described a two-stage implementation: first, proven psychophysical stimuli were integrated into contextual forms, or codex blocks, and second, adapted to the environment of complex visual scenes. This approach argues for direct application in generating new models for learning, interaction, navigation, and begin to define emergent potentials in immersive environments.

Systematic and applied longitudinal studies across a population of subjects yielded significant preliminary findings. The outcomes of experimentation validate the framework and methods for delivering semantic information to far peripheral regions of the human retina as a valuable extension of contemporary methodologies. Outcomes demonstrated by conducted experimentation can be summarized as:

- accurate far peripheral symbol discrimination is achievable within small (< 1 second) timescales

- trends in detection speed over successive trials support early learning adoption of a new visual codex

- far peripheral symbol recognition in highly dynamic visual environments is possible absent of central fixation, fixed contrast, luminance and illuminance

- improvements in detection rates achieved through visual learning in the control series are present in low- and high-load series at varying offsets.

Each codex block was quantified in terms of its absolute motion energy and weighed against rates for accurate symbol detection across all subjects in the control series. Trends in detection time versus symbol complexity were found with notable exceptions – these outliers were analyzed and found to exhibit unique characteristics stemming from their design. Overall, symbol uniqueness within the codex was a strong determining factor for detection speed while occupying comparatively little spatial real estate in both pixel space and retinotopic projections onto discrete regions of the retina.

Further parameterization of these methodologies is a subject of immediate future work. Completed experimentation and its outcomes have contributed toward first principles and establishing feasibility for deployment in real-world environments. I look forward to future study designs that isolate and control for mechanisms of adaptation. In addition, correlating symbol detection rates with codex size in future work can help to elucidate the role of symbol uniqueness in rates of discrimination.

This dissertation argued a new intersection in the fields of vision science, computational imaging, and display technologies. As the technological cutting edge outpaces our physiological sensitivities (in resolution, frame rate, FOV), this work may function as a step in mapping a new generation of biologically encoded systems and potentially offloading some of the computational-media-landscape to early mechanisms underlying the human visual pipeline.

Apart from the immediacy of application across industry in existing technological infrastructures, there is a greater goal to leverage: vision and the extraordinary potentials inherent in sensory mechanisms to challenge society and future technologies toward a better understanding of ourselves. The small steps in this dissertation are by no means a conclusion but are an exciting trajectory to dedicate my energies in earnest. The way to move beyond ourselves is to look further within.

"Accordingly, the first thing we have to learn is to pay heed to our individual sensations. Ordinarily we do so merely in case of those sensations that enable us to find out about the world around us. In the ordinary affairs of life the sensations have no other importance for us. Subjective sensations are of interest chiefly for scientific investigations only. If they happen to be noticed in the ordinary activity of the senses, they merely distract the attention. Thus while we may attain an extraordinary degree of delicacy and precision in the objective observation, we not only fail to do so in subjective observations, but indeed we acquire the faculty in large measure of overlooking them and of forming our opinions of objects independently of them, even when they are so pronounced that they might easily be noticed."

Hermann Ferdinand Ludwig von Helmholtz,

- Treatise on Physiological Optics

9. References

[1]   R. L. Achtman, R. F. Hess, and Y. Z. Wang, "Regional sensitivity for shape discrimination," *Spat. Vis.*, vol. 13, no. 4, pp. 377–391, Jan. 2000, doi: 10.1163/156856800741261.

[2]   J. Schwiegerling, *Field guide to visual and ophthalmic optics*. Bellingham, WA: SPIE, 2004.

[3]   Z. Lu and B. Dosher, *Visual psychophysics: from laboratory to theory*. Cambridge, MA: The MIT Press, 2014.

[4]   C. A. Curcio, K. R. Sloan, R. E. Kalina, and A. E. Hendrickson, "Human photoreceptor topography," *J. Comp. Neurol.*, vol. 292, no. 4, pp. 497–523, Feb. 1990, doi: 10.1002/cne.902920402.

[5]   H. Strasburger, I. Rentschler, and M. Jüttner, "Peripheral vision and pattern recognition: A review," vol. 11, no. 13, pp. 1–82, 2011, doi: 10.1167/11.5.13.Contents.

[6]   J. Rovamo, L. Leinonen, P. Laurinen, and V. Virsu, "Temporal integration and contrast sensitivity in foveal and peripheral vision," *Perception*, vol. 13, no. 6, pp. 665–674, 1984, doi: 10.1068/p130665.

[7]   H. Strasburger, L. O. Harvey, and I. Rentschler, "Contrast thresholds for identification of numeric characters in direct and eccentric view," *Percept. Psychophys.*, vol. 49, no. 6, pp. 495–508, 1991, doi: 10.3758/BF03212183.

[8]   H. Strasburger, I. Rentschler, and L. O. Harvey, "Cortical Magnification Theory Fails to Predict Visual Recognition," *Eur. J. Neurosci.*, vol. 6, no. 10, pp. 1583–1588, 1994, doi: 10.1111/j.1460-9568.1994.tb00548.x.

[9]   A. Shapiro, E. Knight, Z. L.-Pl. One, and undefined 2011, "A first-and second-order motion energy analysis of peripheral motion illusions leads to further evidence of 'feature blur' in peripheral vision," *ncbi.nlm.nih.gov*, Accessed: Feb. 04, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3084698/.

[10]   S. E. Palmer, *Vision science: Photons to phenomenology*. Cambridge, MA: MIT press, 1999.

[11]   L. Ungerleider, J. H.-C. opinion in neurobiology, and undefined 1994, "'What'and 'where'in the human brain," *Elsevier*, Accessed: Mar. 05, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0959438894900663.

[12]   J. D. Victor, K. Purpura, E. Katz, and B. Mao, "Population encoding of spatial frequency, orientation, and color in macaque V1," *J. Neurophysiol.*, vol. 72, no. 5, pp. 2151–2166, 1994, doi: 10.1152/jn.1994.72.5.2151.

[13]   S. O. Dumoulin and B. A. Wandell, "Population receptive field estimates in human visual cortex," *Neuroimage*, vol. 39, no. 2, pp. 647–660, Jan. 2008, doi: 10.1016/j.neuroimage.2007.09.034.

[14]   J. Gibson, "The senses considered as perceptual systems.," 1966, Accessed: Feb. 06, 2020. [Online]. Available: https://psycnet.apa.org/record/1966-35026-000.

[15]   E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," *J. Opt. Soc. Am. A*, vol. 2, no. 2, 1985, Accessed: Aug. 29, 2017. [Online]. Available: https://www.osapublishing.org/DirectPDFAccess/D1C2C850-AE0B-00F9-FFDE5821E0AE32E0_1945/josaa-2-2-284.pdf?da=1&id=1945&seq=0&mobile=no.

[16]   E. H. Adelson and J. R. Bergen, "The Plenoptic Function and the Elements of Early Vision," pp. 3–20, 1991, Accessed: Dec. 04, 2017. [Online]. Available: http://faculty.cs.tamu.edu/jchai/CPSC641/elements91.pdf.

[17] C. Chubb and G. Sperling, "Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception," *J. Opt. Soc. Am. A*, vol. 5, no. 11, p. 1986, 1988, doi: 10.1364/josaa.5.001986.

[18] D. Marr and S. Ullman, "Directional selectivity and its use in early visual processing," *Proc. R. Soc. London B*, vol. 211, no. 1183, pp. 151–180, 1981.

[19] M. Schneider, I. Marquardt, S. Sengupta, F. De Martino, and R. Goebel, "Motion displaces population receptive fields in the direction opposite to motion," 2019, doi: 10.1101/759183.

[20] R. Fredericksen, P. Bex, F. V.-J. A, and undefined 1997, "How big is a Gabor patch, and why should we care?," *osapublishing.org*, Accessed: Jan. 21, 2019. [Online]. Available: https://www.osapublishing.org/abstract.cfm?uri=josaa-14-1-1.

[21] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *J. Neurophysiol.*, vol. 58, no. 6, pp. 1233–1258, 1987, [Online]. Available: https://www.physiology.org/doi/pdf/10.1152/jn.1987.58.6.1233.

[22] P. Bloomfield, *Fourier analysis of time series: an introduction.* 2004.

[23] K. Amano, M. Edwards, D. R. Badcock, and S. Nishida, "Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus," *J. Vis.*, vol. 9, no. 3, pp. 1–25, 2009, doi: 10.1167/9.3.4.

[24] A. Rider, P. McOwan, A. J.-J. of Vision, and undefined 2009, "Motion-induced position shifts in global dynamic Gabor arrays," *jov.arvojournals.org*, Accessed: Feb. 04, 2020. [Online]. Available: https://jov.arvojournals.org/article.aspx?articleid=2193407.

[25] S. M. Thurman and H. Lu, "Bayesian integration of position and orientation cues in perception of biological and non-biological forms," *Front. Hum. Neurosci.*, vol. 8, no. 1 FEB, pp. 1–13, 2014, doi: 10.3389/fnhum.2014.00091.

[26] R. L. De Valois and K. K. De Valois, "VERNIER ACUITY WITH STATIONARY MOVING GABORS," 1991.

[27] D. H. Arnold, M. Thompson, and A. Johnston, "Motion and position coding," *Vision Res.*, vol. 47, no. 18, pp. 2403–2410, Aug. 2007, doi: 10.1016/J.VISRES.2007.04.025.

[28] D. Whitney and P. Cavanagh, "Motion distorts visual space: Shifting the perceived position of remote stationary objects," *Nat. Neurosci.*, 2000, doi: 10.1038/78878.

[29] D. W. Bressler and D. Whitney, "Second-order motion shifts perceived position," *Vision Res.*, vol. 46, no. 6–7, pp. 1120–1128, 2006, doi: 10.1016/j.visres.2005.10.012.

[30] V. S. Ramachandran and S. M. Anstis, "Illusory displacement of equiluminous kinetic edges.," *Perception*, 1990, doi: 10.1068/p190611.

[31] R. Hisakata, D. Hayashi, I. M.-J. of vision, and undefined 2016, "Motion-induced position shift in stereoscopic and dichoptic viewing," *jov.arvojournals.org*, Accessed: Feb. 04, 2020. [Online]. Available: https://jov.arvojournals.org/article.aspx?articleid=2572580.

[32] S. Chung, S. Patel, H. Bedell, O. Y.-V. Research, and U. 2007, "Spatial and temporal properties of the illusory motion-induced position shift for drifting stimuli," *Elsevier*, vol. 47, pp. 231–243, 2007, Accessed: Feb. 04, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0042698906004998.

[33] P. V. McGraw, D. Whitaker, J. Skillen, and S. T. L. Chung, "Motion adaptation distorts perceived visual position," *Curr. Biol.*, vol. 12, no. 23, pp. 2042–2047, 2002, doi: 10.1016/S0960-9822(02)01354-4.

[34] W. R.-S. communication and undefined 1961, "Autocorrelation, a principle for the

evaluation of sensory information by the central nervous system," *ci.nii.ac.jp*, Accessed: Mar. 05, 2020. [Online]. Available: https://ci.nii.ac.jp/naid/10006533604/.

[35]   M. Carrasco, "Visual attention: The past 25 years," *Vision Research*. 2011, doi: 10.1016/j.visres.2011.04.012.

[36]   N. Lavie and Y. Tsal, "Perceptual load as a major determinant of the locus of selection in visual attention," *Percept. Psychophys.*, vol. 56, no. 2, pp. 183–197, Mar. 1994, doi: 10.3758/BF03213897.

[37]   N. Lavie, "Perceptual load as a necessary condition for selective attention," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 21, no. 3, pp. 451–468, 1995, doi: 10.3758/BF03213897.

[38]   A. Mack, Z. Pappas, M. Silverman, and R. Gay, "What we see: Inattention and the capture of attention by meaning," in *Consciousness and Cognition*, 2002, doi: 10.1016/S1053-8100(02)00028-4.

[39]   G. Rees, C. Russell, C. D. Frith, and J. Driver, "Inattentional blindness versus inattentional amnesia for fixated but ignored words," *Science (80-.).*, 1999, doi: 10.1126/science.286.5449.2504.

[40]   H. H. Haladjian, M. Lisi, and P. Cavanagh, "Motion and position shifts induced by the double-drift stimulus are unaffected by attentional load," *Attention, Perception, Psychophys.*, vol. 80, no. 4, 2018, doi: 10.3758/s13414-018-1492-0.

[41]   "Schäffel's 'Luminance Looming.'" https://michaelbach.de/ot/sze-lumiloom/index.html (accessed Jul. 31, 2020).

[42]   O. Kwon, D. Tadin, D. K.-P. of the National, and undefined 2015, "Unifying account of visual motion and position perception," *Natl. Acad Sci.*, Accessed: Feb. 04, 2020. [Online]. Available: https://www.pnas.org/content/112/26/8142.short.

[43]   S. T. L. Chung, S. S. Patel, H. E. Bedell, and O. Yilmaz, "Spatial and temporal properties of the illusory motion-induced position shift for drifting stimuli," *Vision Res.*, vol. 47, no. 2, pp. 231–243, Jan. 2007, doi: 10.1016/j.visres.2006.10.008.

[44]   H. Öğmen and B. Breitmeyer, *The first half second: The microgenesis and temporal dynamics of unconscious and conscious visual processes*. 2006.

[45]   S. Nakamura and S. Takahashi, "An illusory contour can facilitate visually induced self-motion perception," *Multisens. Res.*, vol. 31, no. 8, pp. 715–727, 2018, doi: 10.1163/22134808-20181312.

[46]   I. P. Howard and A. Howard, "Vection: the contributions of absolute and relative visual motion.," *Perception*, vol. 23, no. 7, pp. 745–51, 1994, doi: 10.1068/p230745.

[47]   L. L. Vallerie, "Peripheral vision displays," 1968.

[48]   D. Hameluck and P. Stager, "Instrument Scanning and Subjective Workload with the Peripheral Vision Horizon Display," *Proc. Hum. Factors Soc. Annu. Meet.*, vol. 33, no. 2, pp. 18–22, Oct. 1989, doi: 10.1177/154193128903300204.

[49]   D. J. C. Matthies, M. Haescher, R. Alm, and B. Urban, "Properties of a peripheral head-mounted display (PHMD)," in *Communications in Computer and Information Science*, Aug. 2015, vol. 528, pp. 208–213, doi: 10.1007/978-3-319-21380-4_37.

[50]   E. Costanza, S. A. Inverso, E. Pavlov, R. Allen, and P. Maes, "Eye-q: Eyeglass peripheral display for subtle intimate notifications," in *ACM International Conference Proceeding Series*, 2006, vol. 159, pp. 211–218, doi: 10.1145/1152215.1152261.

[51]   H. Reiterer and & O. Deussen, "Mensch & Computer 2012: 12. fachübergreifende Konferenz für interaktive und kooperative Medien.," Oldenbourg Verlag, 2012.

[52]   I. E. Andes, K. W. Fyfe, and J. P. Mcevoy, "PERPHERAL VISION HEAD-MOUNTED DISPLAY FOR IMPARTING INFORMATION TO A USER WITHOUT

DISTRACTION AND ASSOCATED METHODS," US 9,645,396 B2, 2017.

[53]   T. Nakao, M. Nakatani, L. Chan, and K. Kunze, "Smart Glasses with a Peripheral Vision Display," *dl.acm.org*, Mar. 2016, doi: 10.1145/2927929.2927938.

[54]   P. Lubos, G. Bruder, O. Ariza, and F. Steinicke, "Ambiculus: LED-based Low-Resolution Peripheral Display Extension for Immersive Head-Mounted Displays," doi: 10.1145/2927929.2927939.

[55]   U. Gruenefeld, T. C. Stratmann, A. El Ali, S. Boll, and W. Heuten, "Radiallight: Exploring radial peripheral LEDs for directional cues in head-mounted displays," in *MobileHCI 2018 - Beyond Mobile: The Next 20 Years - 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, Conference Proceedings*, Sep. 2018, vol. 18, pp. 1–6, doi: 10.1145/3229434.3229437.

[56]   U. Gruenefeld *et al.*, "Guiding Smombies: Augmenting Peripheral Vision with Low-Cost Glasses to Shift the Attention of Smartphone Users," in *Adjunct Proceedings - 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2018*, Jul. 2018, pp. 127–131, doi: 10.1109/ISMAR-Adjunct.2018.00050.

[57]   A. Lucero and A. Vetek, "NotifEye: Using interactive glasses to deal with notifications while walking in public," in *ACM International Conference Proceeding Series*, Nov. 2014, vol. 2014-November, pp. 1–10, doi: 10.1145/2663806.2663824.

[58]   H. Y. Tseng, R. H. Liang, L. Chan, and B. Y. Chen, "LEaD: Utilizing light movement as peripheral visual guidance for scooter navigation," in *MobileHCI 2015 - Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Aug. 2015, pp. 323–326, doi: 10.1145/2785830.2785831.

[59]   K. Luyten, D. Degraen, G. Rovelo, S. Coppers, and D. Vanacken, "Hidden in plain sight: An exploration of a visual language for near-eye out-of-focus displays in the peripheral view," in *Conference on Human Factors in Computing Systems - Proceedings*, May 2016, pp. 487–497, doi: 10.1145/2858036.2858339.

[60]   S. Langlois, "ADAS HMI using peripheral vision," *dl.acm.org*, pp. 74–81, 2013, doi: 10.1145/2516540.2516558.

[61]   T. Van Veen, J. Karjanto, and J. Terken, "Situation awareness in automated vehicles through proximal peripheral light signals," in *AutomotiveUI 2017 - 9th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, Proceedings*, Sep. 2017, pp. 287–292, doi: 10.1145/3122986.3122993.

[62]   "Method and apparatus for vehicle enabled visual augmentation," Oct. 2014.

[63]   A. Matviienko, A. Löcken, A. El Ali, W. Heuten, and S. Boll, "NaviLight: Investigating Ambient Light Displays for Turn-by-Turn Navigation in Cars," *dl.acm.org*, pp. 283–294, Sep. 2016, doi: 10.1145/2935334.2935359.

[64]   A. Löcken, W. Heuten, and S. Boll, "Supporting Lane Change Decisions with Ambient Light," *dl.acm.org*, pp. 204–211, 2015, doi: 10.1145/2799250.2799259.

[65]   F. L. Kooi and M. Mosch, "PERIPHERAL MOTION DISPLAYS: TAPPING THE POTENTIAL OF THE VISUAL PERIPHERY."

[66]   A. Meschtscherjakov, C. Döttlinger, T. Kaiser, and M. Tscheligi, "Chase Lights in the Peripheral View: How the Design of Moving Patterns on an LED Strip Influences the Perception of Speed in an Automotive Context," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Accessed: Aug. 01, 2020. [Online]. Available: https://doi.org/10.1145/3313831.3376203.

[67]   P. Renner and T. Pfeiffer, "Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems," in *2017*

IEEE Symposium on 3D User Interfaces, 3DUI 2017 - Proceedings, Apr. 2017, pp. 186–194, doi: 10.1109/3DUI.2017.7893338.

[68] T. Ohshima, … H. Y.-P. of the I., and undefined 1996, "Gaze-directed adaptive rendering for interacting with virtual space," *ieeexplore.ieee.org*, Accessed: Aug. 01, 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/490517/?casa_token=7y_OUh8EmR8AA AAA:mrNPMxOEBGNmqwBBvrQ6XZOfrd792lARhaq-V80WKqShO6zh0RDVAI0EY1kwWFqRdXe29wF8.

[69] S. Matsuzoe, S. Jiang, M. Ueki, and K. Okabayashi, "Intuitive visualization method for locating off-screen objects inspired by motion perception in peripheral vision," in *ACM International Conference Proceeding Series*, Mar. 2017, pp. 1–4, doi: 10.1145/3041164.3041198.

[70] Y. Ishiguro and J. Rekimoto, "Peripheral Vision Annotation: Noninterference Information Presentation Method for Mobile Augmented Reality," 2010.

[71] X. Sun and A. Varshney, "Investigating perception time in the far peripheral vision for virtual and augmented reality," in *Proceedings - SAP 2018: ACM Symposium on Applied Perception*, Aug. 2018, doi: 10.1145/3225153.3225160.

[72] M. Kassner, W. Patera, and A. Bulling, "Pupil: An Open Source Platform forPervasive Eye Tracking and MobileGaze-based Interaction," *Proc. 2014 ACM Int. Jt. Conf. pervasive ubiquitous Comput. Adjun. Publ.*, pp. 1151–1160, 2014, [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/2638728.2641695?casa_token=S8LknFQlivUAAAA A:YivWGXXKc24OXNrqbtZwCdMmUpD6Jr6XpQlYOwzNNtyThIa1tGZZv9QdBKixI OOYD_wtISXBZBFb.

[73] E. R.-J. of pediatric ophthalmology and strabismus and undefined 1985, "Normal saccadic velocities," *healio.com*, Accessed: Jul. 31, 2020. [Online]. Available: https://www.healio.com/ophthalmology/journals/jpos/1985-1-22-1/%7B783f3060-00c1-41da-9e64-efba75d8278f%7D/normal-saccadic-velocities.

[74] Z. L. Lu and G. Sperling, "The functional architecture of human visual motion perception," *Vision Res.*, vol. 35, no. 19, pp. 2697–2722, 1995, doi: 10.1016/0042-6989(95)00025-U.

[75] J. R.-V. Research and undefined 1994, "Directional anisotropy of motion sensitivity across the visual field," *Elsevier*, Accessed: Jul. 31, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0042698994900078.

[76] G. Taibbi, R. Cromwell, K. Kapoor, … B. G.-S. of, and undefined 2013, "The effect of microgravity on ocular structures and visual function: a review," *Elsevier*, Accessed: Aug. 27, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0039625712000951.

[77] "What is HyperSmooth?," 2020. https://gopro.com/help/articles/block/what-is-hypersmooth (accessed Feb. 06, 2020).

[78] D. E. Parker, B.-L. H. Duh, T. a Furness, J. D. Prothero, E. J. Seibel, and B.-L. H. Duh, "Alleviating motion, simulator, and virtual environmental sickness by presenting visual scene components matched to inner ear vestibular sensations," Jan. 2001. Accessed: Aug. 16, 2019. [Online]. Available: https://patents.google.com/patent/US6497649B2/en.

[79] "Methods of vestibulo-ocular reflex correction in display systems," Oct. 28, 2015.

[80] "System and method for motion sickness minimization using integration of attended and unattended datastreams," Aug. 11, 2011.

[81] "COMPUTER-READABLE MEDIA FOR USING A VIDEO CAPTURE DEVICE TO ALLEVIATE MOTION SICKNESS VIA AN AUGMENTED DISPLAY FOR A PASSENGER," 2932.

[82] C. R. Fetsch, A. H. Turner, G. C. DeAngelis, and D. E. Angelaki, "Dynamic reweighting of visual and vestibular cues during self-motion perception," *J. Neurosci.*, vol. 29, no. 49, pp. 15601–15612, Dec. 2009, doi: 10.1523/JNEUROSCI.2574-09.2009.

[83] P. Scarfe and A. Johnston, "Motion drag induced by global motion Gabor arrays," *J. Vis.*, vol. 10, no. 5, pp. 1–15, 2010, doi: 10.1167/10.5.14.

[84] A. L. F. Lee and H. Lu, "A comparison of global motion perception using a multiple-aperture stimulus," *J. Vis.*, vol. 10, no. 4, pp. 1–16, 2010, doi: 10.1167/10.4.9.

Appendix A: Proprioceptive augmentation

The core output of this dissertation is a codex of perceptual filters, optimized through user studies, for far peripheral symbol recognition in dynamic environments. The following chapter describes in further detail the novel systems engineered and articulated to accomplish a further dimension of perceptual augmentation: observer vection. The block diagrams in Figure 37 summarize the flow of key data elements and their algorithmic combinations to yield augmented environments. Note that the flow diagrams for the two perceptual augmentation strategies utilize the same fundamental methods for raw scene data parsing and raw psychophysical structure adaptation – these two strategies are distinguished by the *motion path* which is used to modulate the adapted carrier signals (shown in red).
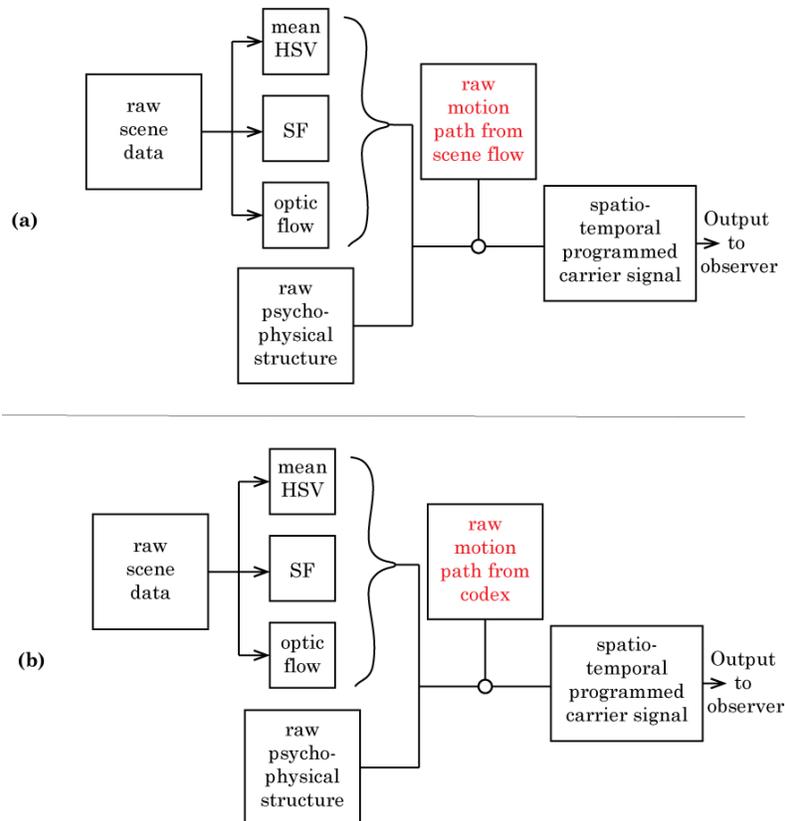
Figure 37. System block diagrams for (a) proprioceptive and (b) semantic sequences. Mean HSV indicates mean hue, saturation, and value; SF indicates spatial frequency.
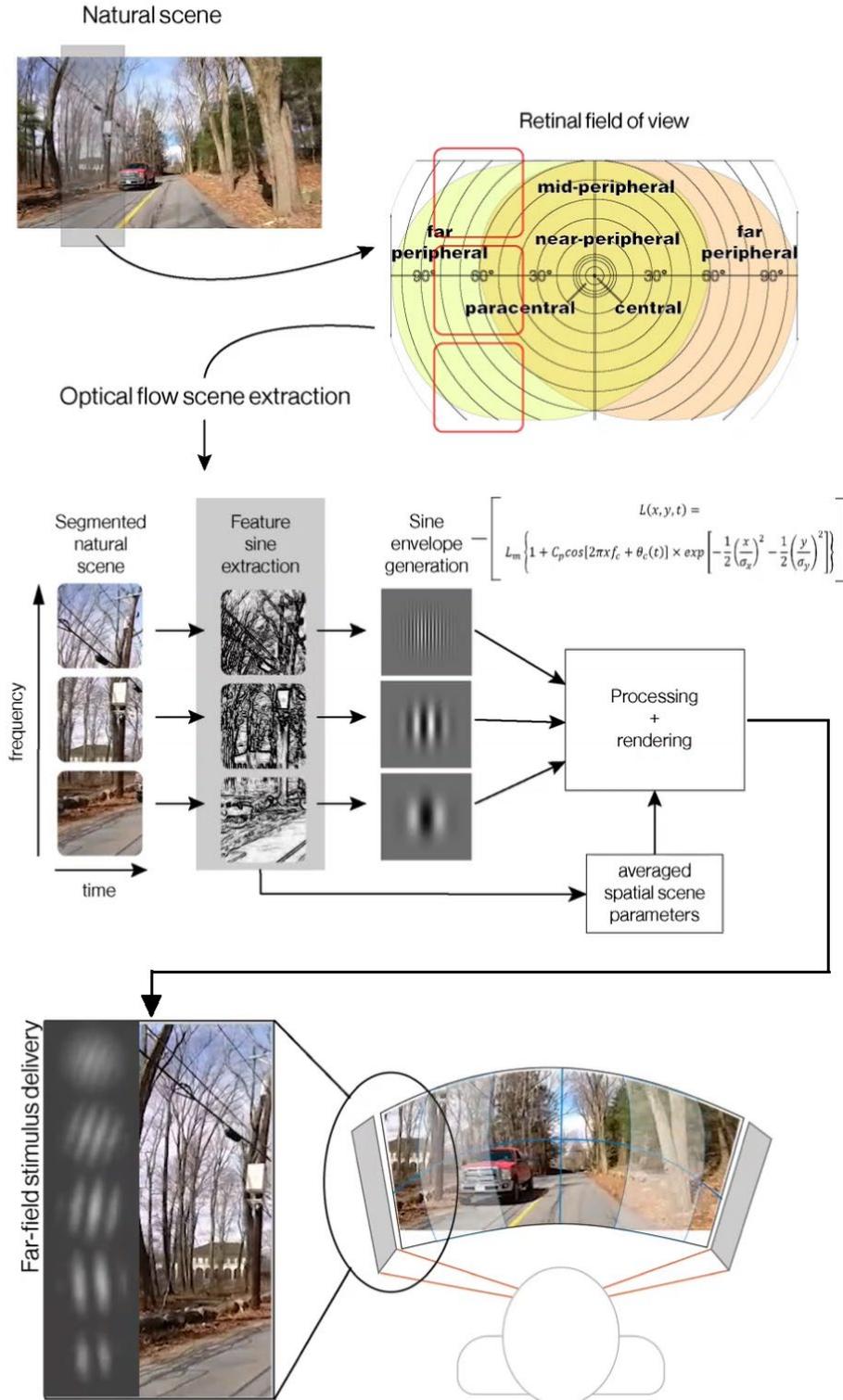
Figure 38. Scene parsing and parameterization for spatial frequency stimulus adaptation. Regions of interest within a natural scene are analyzed for their constituent spatial harmonics. Outputs from spatial frequency analysis adapt raw psychophysical structures to generate contextual stimuli.

Figure 39. Sample still of the display environment (central view) with environmental calibration overlay.

Figure 40. Hardware implementations for proprioceptive (left) and semantic augmentation (right)

While the evaluation of these systems was outside the scope of this dissertation, this work stands as an important extension of the approach at the core of my research – to multiplex perceptual outcomes via computational systems targeting early visual processes. This chapter will first address how we generated stimuli that target and activate discrete visual processes to affect visual perception in dynamic environments, considering both signal adaptation and motion modulation.

The next consideration is of *delivering* these stimuli to the observer such that the features which target discrete visual mechanisms are preserved.

Finally, the system must incorporate methods for observer data *capture* with sufficient bit rate and timing fidelity.

Adaptation and modulation

The novel series of stimuli used throughout the experimentation are generated once the raw environmental footage is stabilized and formatted for the study environment (as detailed in the following section). All raw stimuli are adapted according to the following process:

1. Regions of interest (ROIs) are identified within the environmental footage to serve as raw scene data sources.
2. Frame-by-frame scene parameters of hue, saturation, value, spatial frequency, and optical flow are calculated within the ROIs.
3. Raw psychophysical carrier signals (such as first-and second-order Gabor patches) are generated.
4. Carrier signal motion is determined (either from scene data or codex paths)
5. Adaptive carrier signals are generated from scene data and motion parameters.
6. Carrier signal locations are identified within the simulated environment within the observer field of view.
7. Carrier signal envelopes are generated (either traditional Gaussian windows or scene-adaptive, non-uniform envelopes, measured in arc min)
8. Adaptive carrier signals are windowed according to the generated signal envelopes and integrated at the defined locations within the simulated environment.

In initial evaluations conducted to validate this system throughput, we primarily used sinusoidal gratings and Gaussian white noise as raw psychophysical carrier signals, generated in MATLAB 2019b. Figure 41 shows parametric variations of two standard psychophysical structures ("sinusoidal grating" and "white Gaussian noise"), as well as Gaussian windowed scene samples ("raw scene data"), demonstrating some of the spans of programmable stimuli used in this systems approach. An important avenue of future work is the quantification of perceptual outcomes from using these three varied stimuli.
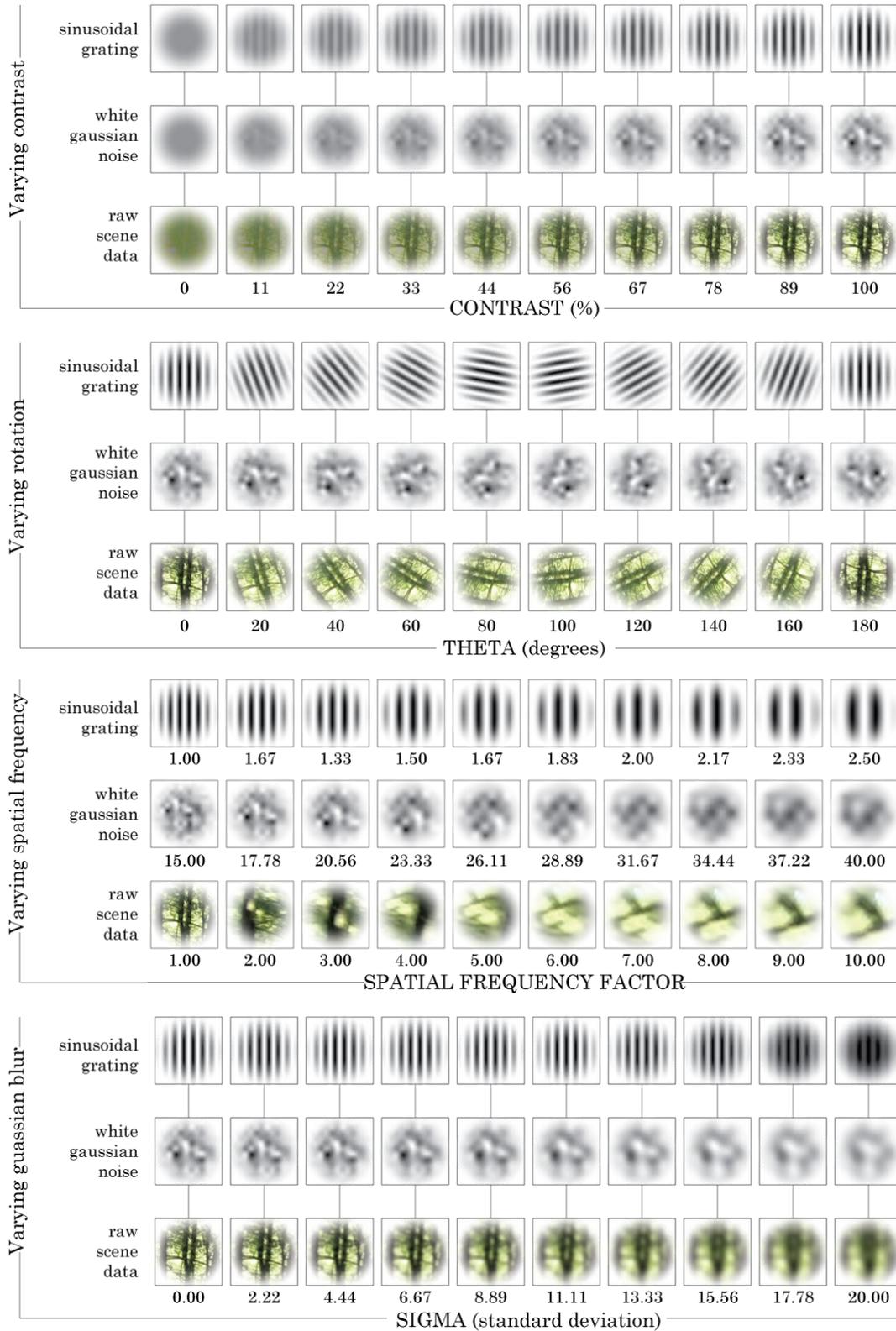
Figure 41. Our generation of psychophysical structures and natural scene data varied parametrically with contrast, rotation, spatial frequency and Gaussian blur. [include measurement of cycles per degree visual angle for sinusoidal gratings]

<u>Chromatic adaptation</u>

As in the case of implementing the codex in dynamic natural scenes, methods of chromatic adaptation of stimuli are summarized here.
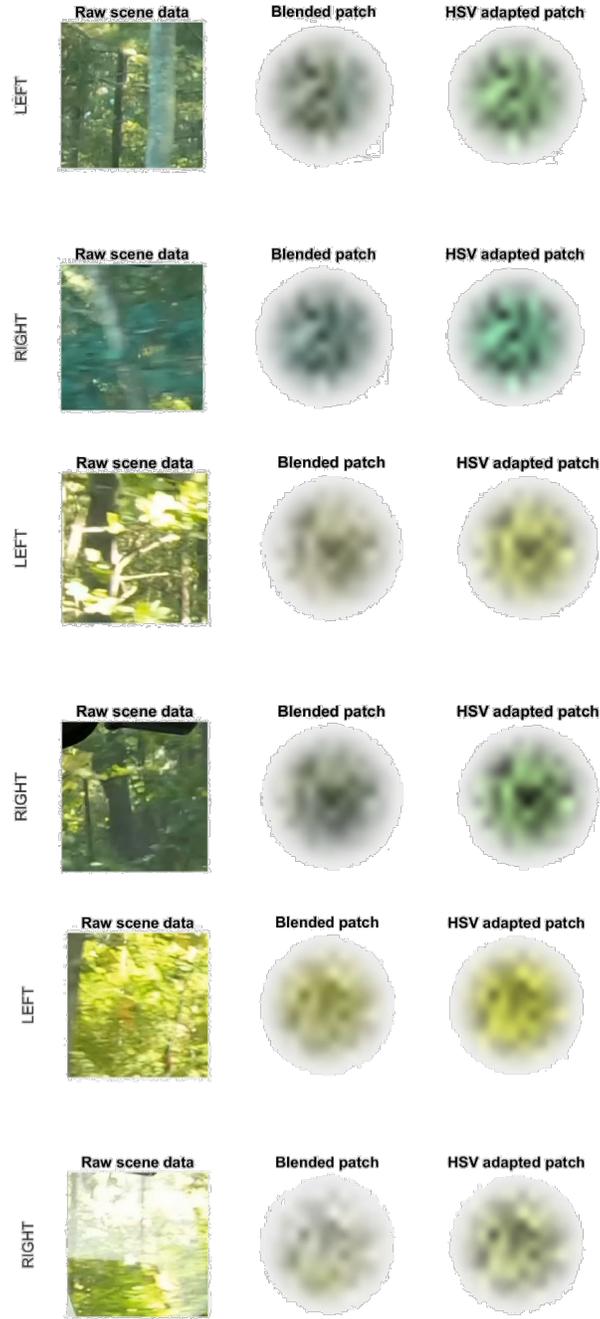
Figure 42. Adapted psychophysical structures, demonstrating two methods of adaptation: overlay blending (middle column), and 1:1 scaled HSV adaptation.

Sample adapted psychophysical structures are shown in Figure 42. The source psychophysical structures are second-order Gabor patches: windowed Gaussian white noise

blurred with a 2-D Gaussian smoothing kernel, σ = 5. Two methods of adaptation are highlighted here – overlay blending and 1:1 hue, saturation, value (HSV) adaptation. The former is a variation of element-wise multiplication of the two source images which results in a linear interpolation of the top layer between the span of values in the base layer:

$$f(a, b) = \begin{cases} 2 * a * b, & 0 \le a < 0.5 \\ 1 - 2 * (1 - a) * (1 - b), & 0.5 \le a \le 1 \end{cases}$$

Alternatively, HSV adaptation transforms the source patch RGB matrix to HSV space and updates the hue and saturation layers to match the mean values from the environmental footage. For the experimentation conducted to support this dissertation, the overlay blending method was used, favored for its subtler integration of stimuli within the environment. Further work is needed to quantify the relative outcomes of using either adaptation methodology.

Spatial frequency adaptation

Using Fourier methods of digesting two-dimensional arrays into their frequency components, the source footage is analyzed to obtain a mean assessment of the feature size within the regions of interest.
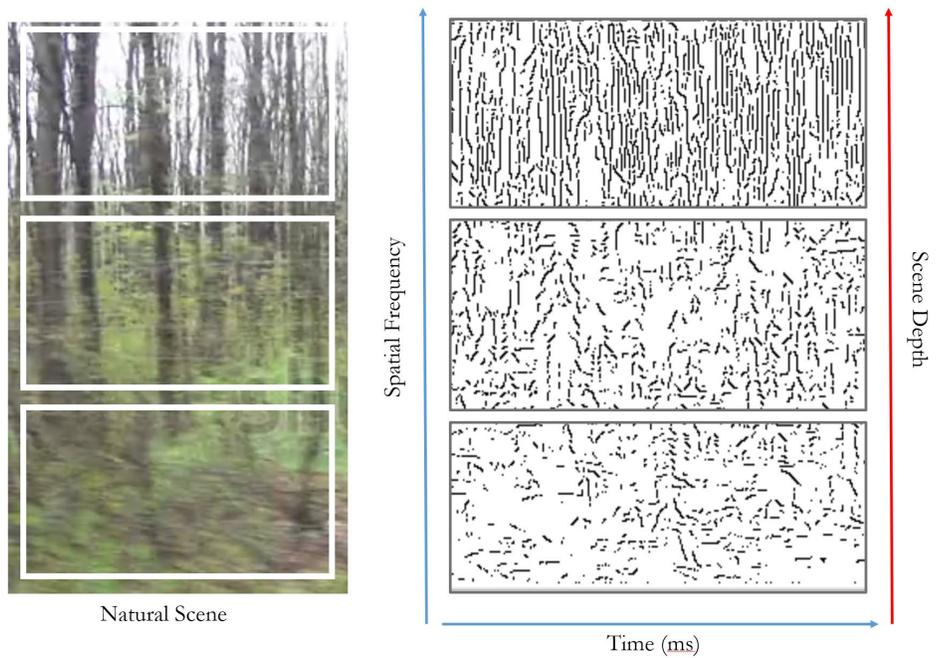
Figure 43. Sample scene parsing. Spatial frequency calculations vary significantly between ROIs near the top of the image (high frequency, distant image features) and the bottom of the image (low frequency, near image features)

Proprioceptive augmentation: motion stimulus from environmental parameters

Once the raw psychophysical structures have been adapted to reflect the chromatic and spatial frequency characteristics of the scene, they are next animated according to motion path data; for the proprioceptive case, this data is derived from scene optical flow calculations. The figure below shows example scene optical flow calculations. The outputs from these calculations determine the *motion* of the stimulus overlay – for example, when optical flow is greater on the righthand side of the scene than the left, that indicates a rightward motion trajectory with its magnitude given by the disparity between the left and right optical flow measurements. By delivering stimuli that, for example, create peripheral motion artifacts proportional to the scene optical flow measurements, this proposed system can either amplify or diminish perceived rate of turn.

Source footage for proprioceptive augmentation was collected in a variety of environments, using capture rigs situated both in the interior and on the exterior of the vehicle. Interior footage is particularly susceptible to artifacts caused by rapid changes in brightness and glare on glass surfaces, which yields significant noise pollution in the optical flow calculations. Specifically, the emergence of stationary glare artifacts corresponds to sharp dips in optical flow data. The most straightforward way to remove such artifacts is to utilize video from the outside of the vehicle (as is the case for the suburban footage used for observer speed perception).
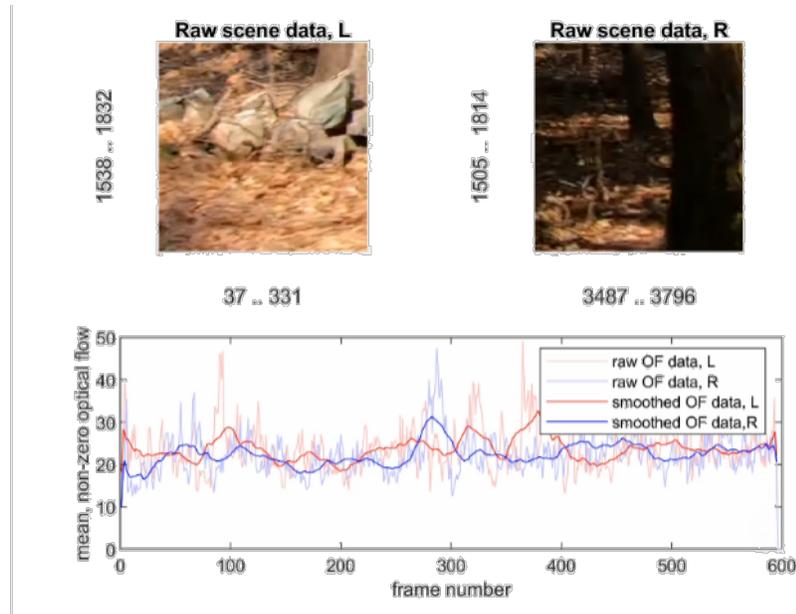
Figure 44. Sample optical flow data, showing the relative relationship of flow calculations from diametrically opposed ROIs. Source footage captured from external POV to minimize artifacts.

As for the other general noise artifacts, a smoothing filter on the data sufficiently captures the mean changes in optical flow which represent the changes in the velocity of the vehicle and can be effectively used to update the psychophysical cues. A moving average with a large spanning window (in this case 30 data points, equal to one second of video footage), while still susceptible to the "dips" caused by glare artifacts, best captures the general trends of the mean optical flow data.

The regions of interest in two footage sources (rural and suburban) were chosen to maximize the fidelity of the optical flow measurements. ROI locations were selected to target surfaces in the scene whose distance from the image sensors varies little throughout footage capture. This approach ensures that changes in optical flow account predominantly for changes in the velocity of the vehicle and not variations in the distance of features from the image sensor. Regions near the ground plane immediately beyond the vehicle (if visible) work best for this approach. Future work in this space will explore algorithmic methodologies for measuring this fidelity and adapting ROIs to maximize fidelity coefficients. For the scope of this dissertation, any instances of significant deterioration of the ROI's representation of scene motion were addressed on a case by case basis.

Curve perception

The first investigation of proprioceptive peripheral feedback concerns the observer's experience of rate of turn, or angular heading. The source footage rate of turn magnitude was determined by a ratio of mean optical flow values from diametrically opposite ROIs.

Rate of turn from stabilization data

One benefit from the system architecture stems from the GoPro HERO7 on-board image stabilization. Broadly, this stabilization functions by slightly cropping the captured footage to account for image jitter, and employs an advanced roller shutter correction algorithm [65]. This yields a rotation of the static elements in the scene (i.e. the vehicle dashboard) around the line of sight axis, mapping directly to the rate of turn. Stabilizing the video motion using the pre-placed markers within the vehicle (see Figure 45) generates position keyframes with high fidelity for predicting rate of turn (see Figure 46 below).
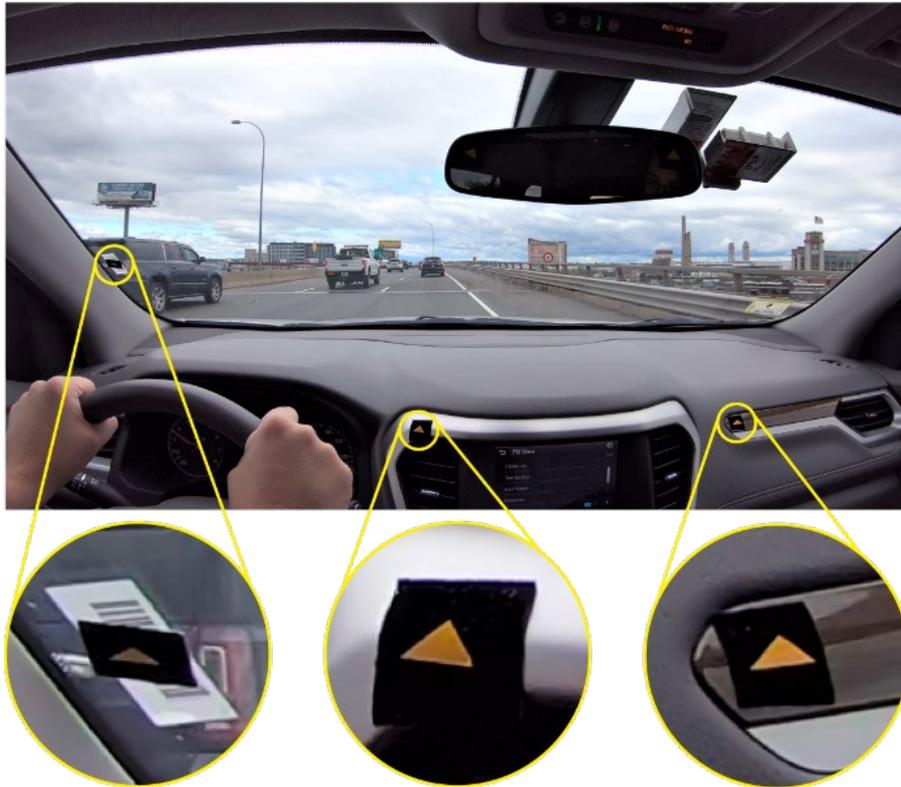
Figure 45. Registration marks for post-processing and stabilization (center view).
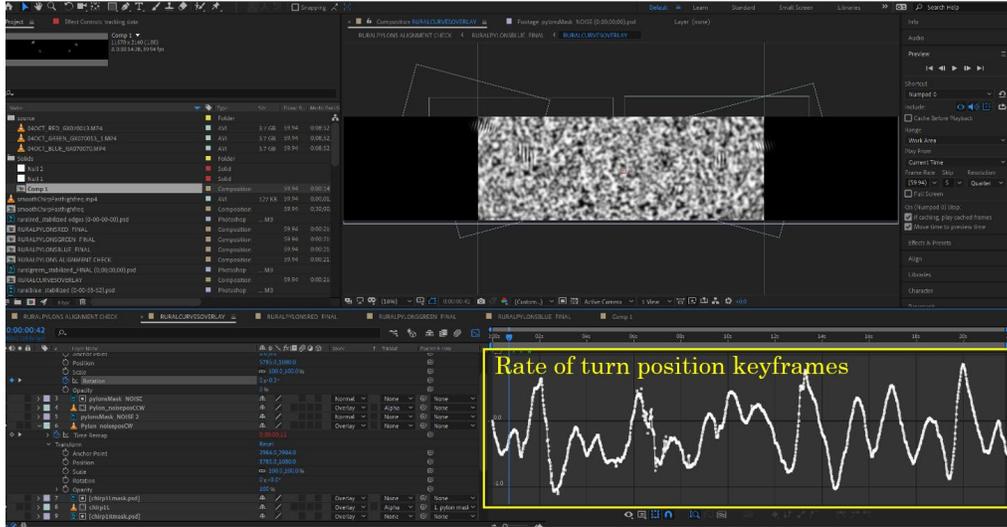
Figure 46. Sample stabilization position keyframes for a "curvy road" sequence, extracted using built-in motion stabilization in Adobe After Effects (indicated in yellow).

Whether using ratios of optical flow in diametrically opposed ROIs or stabilization keyframes, the calculated coefficient for rate of turn was used to scale, frame by frame, the velocity of the carrier signals along their motion trajectories in the periphery. Proprioceptive feedback intended to <u>increase</u> the observer's perception of rate of turn was scaled by the coefficient; proprioceptive feedback intended to <u>decrease</u> the observer's perception of rate of turn was scaled by the inverted coefficient. Positive scalar factor multiplication on the dominant side (in the heading direction) with negative scalar factor multiplication on the non-dominant side (opposite the heading direction resulted in <u>constructive</u> curve augmentation, while the inverse (negative scalar multiplication on the dominant side with positive scalar multiplication on the non-dominant side) resulted in <u>deconstructive</u> curve augmentation.

Speed perception (vection)

The second investigation of proprioceptive peripheral feedback targets the observer's experience of self-motion velocity. In this case, the selection of source ROI for calculating optical flow was much more lenient, only requiring a single window for mean flow speed calculation. As in the rate of turn sequence, calculated coefficients representing the speed of the vehicle were used to scale the peripheral feedback stimuli: unaltered coefficients were used to increase and inverted coefficients were used to decrease the observer's perception of self-motion velocity, with carrier signal speeds on the left/right sides of the vehicle interior equivalent.

The following images (Figure 47) demonstrate a complete processing pipeline (from top to bottom).

(A) Raw scene data

(B) Raw psychophysical structures

(C) Target scene locations for stimulus delivery

(D) Overlaid psychophysical structures with target locations demarcated

(E) Still frame of rendered environment
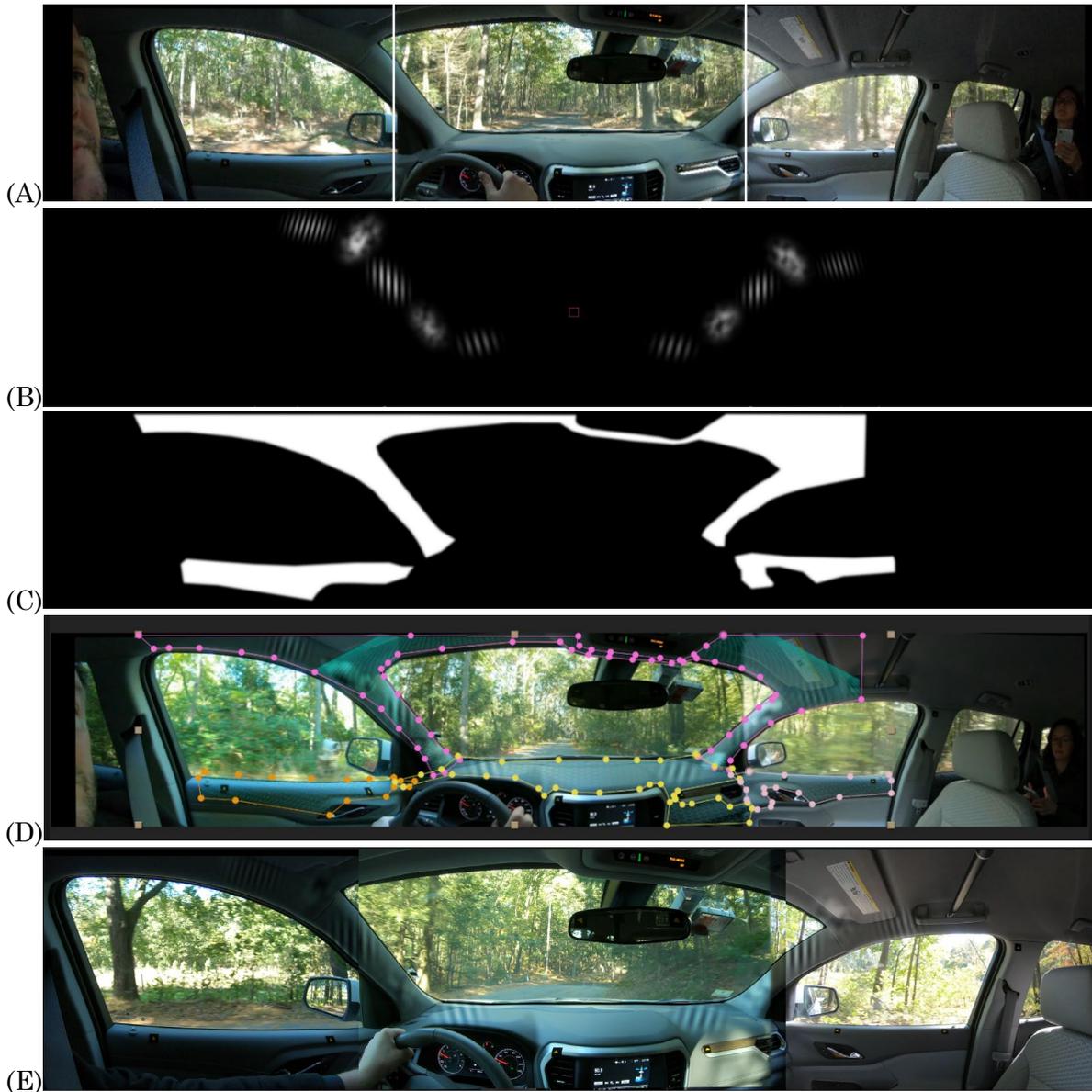
Figure 47. Sample pipeline of integration of proprioceptive feedback stimuli within the vehicle interior.

Footage preparation

Internal vs. external footage capture

In addition to footage capture from the vehicle interior, I designed and implemented the system shown in Figure 48 for external data capture. External footage capture is useful for many testing purposes, being less susceptible to artifacts from internal reflections and shadows. However, it is ultimately less impactful as a study medium as it does not provide static features within the raw footage to serve as projection surfaces for delivering the adapted stimuli.

Figure 48. Externally mounted capture and processing system for raw scene data acquisition and initial segmentation.

Figure 49. Simulated environment, three channel views, **pre-alignment** (rural environment)

Appendix A: Proprioceptive augmentation

Figure 50. Simulated environment, three channel views, **post-alignment** and corner pin effect (rural environment)

Even with onboard image-stabilization, the raw videos contain significant jitter. Before capturing the videos, we placed registration marks in several locations within the interior of the vehicle and afterward imported the raw videos into After Effects to track the motion of the marks and stabilize the footage. After stabilization, the three views were aligned using a corner pin effect and the results were rendered and encoded with the H.265 HEVC codec.

Suburban footage capture

The suburban driving footage was captured by leveling three 4k wide-angle cameras, anchored to the hood of the vehicle with sufficient overlap in field of view to allow for alignment in post-processing. This footage capture configuration was designed for evaluating stimuli without the presence of a static structural element (such as the dashboard) within the scene.

Figure 51. Simulated environment, three channel views, **pre-alignment** (suburban environment)

Figure 52. Simulated environment, three channel views, **post-alignment** and corner pin effect (suburban environment)

Pedestrian footage capture

Figure 53. Raw source footage for pedestrian sequence

Figure 54. Tracked screen mask with peripheral cues, pedestrian sequence

To remove reflections on the phone screen and prepare that region for cues to the observer, the footage had to be first stabilized and then motion tracked to align the screen mask with the movement of the phone throughout the footage.

The tracking algorithms work best when they have small, high contrast features that are consistent throughout the footage (unaffected by changing shadows, color balance, etc.). Maxing out the contrast and brightness, then applying an edge detection filter, improves the speed and accuracy of this process dramatically. Even with these improvements, however, it takes several hours of fine-tuning the parameters to achieve consistent, usable results. Once this process is complete, the tracking data can be applied to either stabilize the footage or animate a separate object.

Generally, when tracking footage in order to mask an object in the scene, there are two options: 1) track the object and apply the tracking data to the mask, or 2) stabilize the object and leave the mask stationary. In the rural driving sequence, for example, the goal was to align the car interior at the seams between the three videos, so the footage had to be stabilized and the interior features anchored to stay aligned throughout the footage. In the pedestrian case, the first attempt was to simply track the phone screen and apply that data to the mask.

Appendix A: Proprioceptive augmentation
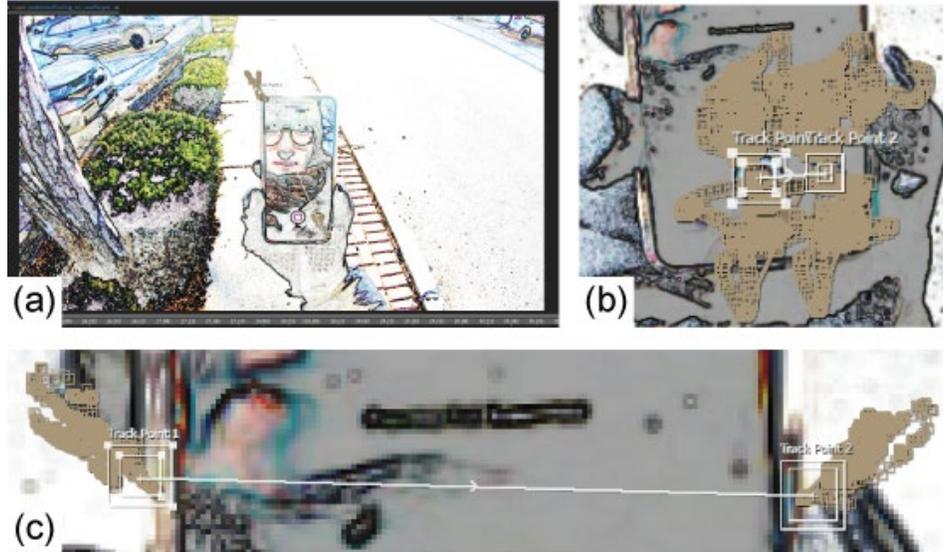
Figure 55. Filtered scene (a), sample first-round tracking points for stabilization (b), and secondary tracking data for occlusion positioning (c)

After troubleshooting the inaccurate tracking outputs, the best approach was to first stabilize the *position* and *rotation* of the phone screen, then track any remaining position and rotation movements not captured in the first tracking step and apply the results of the second tracking step to the mask. As the pictures above show, the sample tracking points for the first round capture a wide range of motion, while the tracking points from the second round capture the remaining motion, which is significantly more contained.

Appendix B: Supplemental results

Gaze dispersion

Gaze tracking was implemented throughout the studies using a Pupil Core headset with a sampling rate at 200Hz and 100° scene field of view, yielding 3D gaze points through binocular vergence. A fixation detection algorithm based on protocols developed by Salvucci and Goldberg[11] was developed in MATLAB R2019b and applied to all subject gaze data output by the Pupil Core headset to identify any instances of gaze diversion exceeding 30 degrees from central fixation. An example of the outputs of our algorithm is shown below.

Symbol	Mean dispersion (degrees)	Variance (σ^2)
C	0.27	0.0361

---
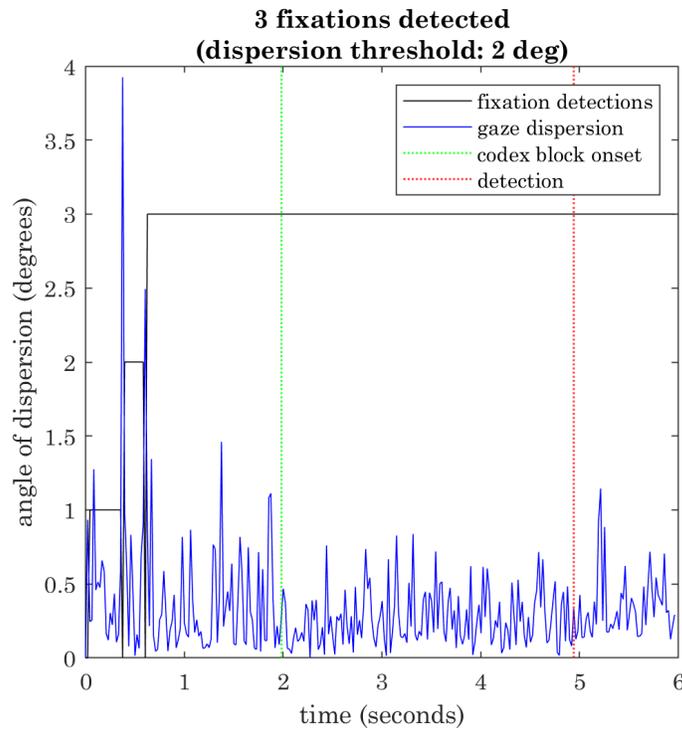
[11] Salvucci, Dario D., and Joseph H. Goldberg. "Identifying fixations and saccades in eye-tracking protocols." Proceedings of the 2000 symposium on Eye tracking research & applications. 2000.

Motion energy analysis

Results from motion energy analysis for the complete codex are provided here. The parts of each graphic are as follows: (A) Single frame excerpt of stimulus video with cross-section annotations indicating locations of x-t projections (red) and y-t projections (blue). (B, C) Projections in x-t and y-t space at cross-sections indicated in (A). Each column represents a fixed aperture in space. The absence of motion in a given aperture is represented by repeating vertical pixels. (D, E) Fourier analysis of motion energy in fx-ft and fy-ft planes with aggregate motion energy value indicated below. Harmonics from the associated projection are indicated by corresponding color (x – red, y – blue).

Please refer to section "7.1 Symbol detection rates correlate with motion energy with notable exceptions" for a generalized interpretation of these results and how they correlate to metrics of codex block perceptibility.

MEA_sym__boatv2
duration = 3.39 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = 2.34$

$E_y = 3.07$

Appendix B: Supplemental results

MEA_sym__flowerv2
duration = 6.69 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = 2.93$

$E_y = 4.48$

MEA_sym__stickfigurev2
duration = 5.06 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = 7.81$

$E_y = -2.00$

MEA_sym__treev2
duration = 4.09 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = -2.31$

$E_y = 2.15$

MEA_sym_2v2
duration = 1.57 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = 1.74$

$E_y = -0.15$

MEA_sym_3v2
duration = 1.70 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = -0.04$

$E_y = 0.90$

MEA_sym_4v2
duration = 1.30 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = -1.24$

$E_y = 0.35$

MEA_sym_5v2
duration = 2.04 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = -0.02$

$E_y = 0.58$

MEA_sym_Av2
duration = 2.04 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = -1.75$

$E_y = -1.18$

Appendix B: Supplemental results

MEA_sym_Bv2
duration = 2.79 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

$E_x = 1.87$

y-t projection harmonics

$E_y = 0.43$

MEA_sym_Cv2
duration = 1.00 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = 0.55$

$E_y = 1.18$

Appendix B: Supplemental results

MEA_sym_Dv2
duration = 2.17 seconds

x-t projection (red)

y-t projection (blue)

x-t projection harmonics

y-t projection harmonics

$E_x = -0.47$

$E_y = 1.75$

<u>Control series</u>

The following section contains the complete raw data outputs for detection speed from each subject in the control series. For each graph, the presented symbol is indicated on the horizontal axis with detection time in seconds on the vertical. Absent data points indicate that the observer response failed to meet the criteria for accurate symbol recognition. Those criteria are:

1. The codex block is accurately identified.
2. Detection was indicated by button press while the codex block was present.
3. Fixation did not deviate in excess of 10 degrees from central fixation cue (in control series only).

This data is arranged longitudinally in a single graphic for each observer in Figure 31, along with fitted trends. For full analysis of trends in detection speed over successive trials, please see section "7.3 Near-minimal detection speeds are achievable for the majority of subjects."

Subject 01, control

Trial 01: $\mu = 5.67$, $\sigma = 1.60$

Trial 02: $\mu = 5.53$, $\sigma = 1.59$

Trial 03: $\mu = 4.21$, $\sigma = 1.76$

Trial 04: $\mu = 3.76$, $\sigma = 1.71$

Trial 05: $\mu = 4.32$, $\sigma = 2.34$

Trial 06: $\mu = 4.79$, $\sigma = 2.49$

Trial 07: $\mu = 2.63$, $\sigma = 1.27$

Trial 08: $\mu = 3.42$, $\sigma = 1.98$

Trial 09: $\mu = 3.15$, $\sigma = 1.91$

Trial 10: $\mu = 3.16$, $\sigma = 1.92$

Subject 02, control

Trial 01: $\mu = 5.60$, $\sigma = 2.02$

B FL BT C FL 5 SF 4 A 2 3 D SF 2 C TR BT D 3 B 5 A 4 TR

Trial 02: $\mu = 5.51$, $\sigma = 2.46$

BT 4 4 SF 2 FL A 2 3 B C BT SF TR 5 FL D D 5 3 TR A B C

Trial 03: $\mu = 3.71$, $\sigma = 1.03$

SF 2 TR A 3 C SF FL 2 4 TR B B A 4 FL 5 D D BT 3 BT C 5

Trial 04: $\mu = 3.91$, $\sigma = 1.44$

A D A 5 2 TR B D FL SF BT 4 C 3 2 B 5 4 3 FL SF TR C BT

Trial 05: $\mu = 3.00$, $\sigma = 1.35$

C 4 SF D 2 FL SF D TR C 3 B A TR 3 FL A B 5 BT 4 5 BT 2

Trial 06: $\mu = 3.07$, $\sigma = 1.21$

FL 5 SF C D B A TR 4 A C FL 3 BT B 3 2 2 D TR BT SF 4 5

Trial 07: $\mu = 2.85$, $\sigma = 0.98$

FL 3 SF 4 B C C B 5 D SF BT D FL 5 TR A BT 2 2 3 4 A TR

Trial 08: $\mu = 2.19$, $\sigma = 0.76$

SF C D SF FL BT 3 TR 2 4 A 5 B A D BT TR B C 4 3 5 2 FL

Trial 09: $\mu = 2.15$, $\sigma = 0.77$

5 5 SF 2 2 TR A 4 3 SF A TR FL BT D C D FL 4 B B C BT 3

Trial 10: $\mu = 2.29$, $\sigma = 0.76$

B A BT SF BT 4 5 3 FL 4 B 3 2 D SF A C 2 TR 5 C FL TR D

Subject 03, control

Trial 01: $\mu = 5.38$, $\sigma = 1.89$

Trial 03: $\mu = 5.47$, $\sigma = 1.82$

Trial 04: $\mu = 5.28$, $\sigma = 1.28$

Trial 05: $\mu = 4.54$, $\sigma = 1.53$

Trial 06: $\mu = 4.43$, $\sigma = 1.79$

Trial 07: $\mu = 4.27$, $\sigma = 2.12$

Trial 08: $\mu = 3.76$, $\sigma = 1.59$

Trial 09: $\mu = 3.83$, $\sigma = 1.76$

Trial 10: $\mu = 4.57$, $\sigma = 1.67$

Subject 04, control

Trial 01: $\mu = 10.33$, $\sigma = 3.80$

FL BT D SF A 5 BT 4 B C 3 4 SF TR D TR 2 C A FL B 5 2 3

Trial 02: $\mu = 9.23$, $\sigma = 2.88$

D FL 4 A FL 5 SF SF D 4 TR 2 BT B 2 B C 3 A BT 5 3 C TR

Trial 03: $\mu = 7.73$, $\sigma = 2.60$

A C 3 SF D FL SF 4 C TR 5 4 B B BT BT TR 2 5 3 2 D FL A

Trial 04: $\mu = 7.28$, $\sigma = 2.76$

SF TR BT B SF B 2 3 C D 4 4 5 TR C 5 FL FL 2 A A D 3 BT

Trial 05: $\mu = 4.79$, $\sigma = 1.02$

5 D FL SF B 4 SF 2 3 A 5 TR FL B C 2 BT TR 4 BT C D 3 A

Trial 06: $\mu = 4.26$, $\sigma = 1.13$

BT B 5 2 3 FL BT D A FL TR C SF 4 D TR C 3 2 4 SF A 5 B

Trial 07: $\mu = 2.45$, $\sigma = 0.87$

3 5 3 D 4 TR TR B D C B A 5 4 SF SF FL 2 2 FL A C BT BT

Trial 08: $\mu = 2.13$, $\sigma = 0.61$

A SF C 4 A D B 4 FL 5 2 BT BT SF 5 D B C TR FL TR 2 3 3

Trial 09: $\mu = 2.24$, $\sigma = 0.65$

3 SF FL 2 D 3 TR BT 5 B A SF TR 4 C FL BT D A 5 C 2 4 B

Trial 10: $\mu = 2.42$, $\sigma = 0.88$

2 C B BT TR B 3 4 TR 4 5 SF A D C FL D FL BT 5 A 2 3 SF

Subject 05, control

Trial 01: $\mu = 2.15$, $\sigma = 1.06$

2 C TR D BT D BT A 5 2 TR 3 3 SF 5 4 FL 4 FL BT C SF BT A

Trial 02: $\mu = 2.12$, $\sigma = 0.76$

5 B TR C 4 B D SF D SF A BT 2 C FL TR A FL 2 5 3 3 BT 4

Trial 03: $\mu = 2.12$, $\sigma = 0.68$

C D C 4 SF FL BT 2 FL 2 A TR B 5 5 4 BT A 3 TR SF B D 3

Trial 04: $\mu = 1.83$, $\sigma = 0.62$

B 3 5 BT 2 2 D TR B C TR C D SF FL A BT 3 SF 4 A 5 4 FL

Trial 05: $\mu = 1.94$, $\sigma = 0.66$

B 5 FL A SF 3 4 2 4 C D 3 SF TR C BT A TR FL D B 5 BT 2

Trial 06: $\mu = 1.80$, $\sigma = 0.53$

FL 2 5 SF 3 2 A B C A 4 TR FL D 5 4 C D BT BT B SF 3 TR

Trial 07: $\mu = 1.70$, $\sigma = 0.51$

3 4 C 2 TR BT 4 FL C 2 BT B SF D TR A B 5 SF 3 FL A D 5

Trial 08: $\mu = 1.89$, $\sigma = 0.53$

2 3 FL D C SF A SF TR TR B A BT 2 3 B D BT 5 5 4 FL 4 C

Trial 09: $\mu = 1.97$, $\sigma = 1.04$

B B SF C C A 4 BT 5 D TR TR SF 2 A 2 D FL 4 FL BT 5 3 3

Trial 10: $\mu = 1.65$, $\sigma = 0.43$

BT D 2 4 TR 3 B 5 5 3 A 4 2 FL SF C SF BT D C TR A B FL

Subject 06, control

Trial 01: μ = 14.49, σ = 4.60

Trial 02: μ = 12.02, σ = 5.23

Trial 03: μ = 8.99, σ = 2.03

Trial 04: μ = 6.41, σ = 1.85

Trial 05: μ = 4.69, σ = 1.49

Trial 06: μ = 2.86, σ = 0.75

Trial 07: μ = 2.29, σ = 0.71

Trial 08: μ = 2.14, σ = 0.78

Trial 09: μ = 2.31, σ = 0.91

Trial 10: μ = 1.94, σ = 0.53

Subject 07, control

Subject 08, control

Trial 01: $\mu = 10.00$, $\sigma = 3.27$

TR TR A D BT 2 B 5 B 4 SF 4 A 2 5 SF BT C FL FL 3 3 D C

Trial 02: $\mu = 4.97$, $\sigma = 0.97$

FL 2 D BT C FL 5 TR SF A TR B 3 4 3 SF 4 2 D B 5 BT C A

Trial 03: $\mu = 3.12$, $\sigma = 0.92$

B B D FL A 5 A BT 3 FL 2 TR 4 SF 3 D SF 5 2 C 4 TR C BT

Trial 04: $\mu = 2.41$, $\sigma = 0.93$

2 TR SF 3 C 5 FL TR A 5 D BT 3 2 SF 4 B A 4 C B BT D FL

Subject 09, control

Trial 01: μ = 14.19, σ = 7.49

FL B BT D SF A 3 B 2 FL 2 4 D 5 4 BT A 5 C 3 TR TR SF C

Trial 02: μ = 11.66, σ = 5.46

4 C BT 3 4 TR 5 5 A B FL SF 2 3 FL TR A SF D C BT B D 2

Trial 03: μ = 8.78, σ = 3.45

FL 3 4 FL A A SF 5 2 5 C D TR B 4 BT SF 3 C BT 2 TR B D

Trial 04: μ = 10.21, σ = 3.88

BT A D SF 5 B 2 SF 2 A 5 TR 3 D 3 4 4 BT TR C FL C FL B

Trial 05: μ = 8.62, σ = 2.20

3 TR A 4 FL BT C 4 2 D A SF C 5 2 BT FL TR B SF B 5 3 D

Trial 06: μ = 8.22, σ = 3.27

3 4 2 FL TR 5 3 B TR B SF BT FL 2 SF C A 4 A BT C D D 5

Trial 07: μ = 9.25, σ = 2.83

TR BT 4 B D A B SF 5 BT 4 2 FL 3 3 TR C 2 FL A 5 SF C D

Trial 08: μ = 6.99, σ = 1.86

D 2 TR A 4 B SF 5 FL 2 D B 3 4 SF 5 3 BT A C TR C FL BT

Trial 09: μ = 6.67, σ = 1.86

A 4 TR 2 2 A 3 FL D C SF 5 BT FL 4 B C D 3 BT TR 5 B SF

Trial 10: μ = 5.92, σ = 1.79

BT BT 5 4 B 5 D 2 A C 4 3 B FL SF SF A TR 2 3 D C TR FL

Subject 10, control

Trial 01: $\mu = 13.76$, $\sigma = 4.66$

TR 4 BT 3 B 5 A D 2 FL BT 4 TR SF 5 B C A 3 SF 2 FL D C

Trial 02: $\mu = 12.79$, $\sigma = 3.03$

2 A 2 D 4 5 4 SF C FL 3 BT 5 TR A B B FL D TR 3 BT SF C

Trial 03: $\mu = 11.92$, $\sigma = 3.76$

C 4 A BT D B 5 TR 2 3 2 A 4 3 B SF FL 5 TR D SF C FL BT

Trial 04: $\mu = 7.94$, $\sigma = 2.31$

D BT C TR C 5 D 5 FL 3 3 A 4 A BT SF 2 TR SF B 2 4 B FL

Trial 05: $\mu = 6.73$, $\sigma = 2.21$

TR B 3 SF BT 4 A SF 3 D D FL 5 A 2 2 C 5 4 B FL C TR BT

Trial 06: $\mu = 5.61$, $\sigma = 2.06$

3 TR SF 3 A 2 5 TR FL BT B BT 2 B D 5 A 4 SF D FL C 4 C

Trial 07: $\mu = 4.74$, $\sigma = 1.18$

D SF BT FL TR C 2 C D B BT 2 A B 4 3 SF 5 5 A 4 TR FL 3

Trial 08: $\mu = 3.75$, $\sigma = 0.82$

B FL SF 2 FL 5 C 4 C A 5 4 D 2 TR B 3 A BT D TR 3 SF BT

Trial 09: $\mu = 2.72$, $\sigma = 0.63$

3 BT TR 2 FL SF 4 D B D 3 TR C FL C A 5 SF B 5 2 A BT 4

Trial 10: $\mu = 2.30$, $\sigma = 0.50$

4 FL 2 C SF BT 4 A 3 BT SF TR 5 FL C B D A B 2 TR D 3 5

Appendix B: Supplemental results

Subject 11, control

Trial 01: $\mu = 2.40$, $\sigma = 0.73$

Trial 02: $\mu = 2.30$, $\sigma = 1.03$

Trial 03: $\mu = 2.05$, $\sigma = 0.60$

Trial 04: $\mu = 1.99$, $\sigma = 0.70$

Trial 05: $\mu = 1.47$, $\sigma = 0.43$

Trial 06: $\mu = 1.57$, $\sigma = 0.57$

Trial 07: $\mu = 1.68$, $\sigma = 0.52$

Trial 09: $\mu = 1.65$, $\sigma = 0.51$

Trial 10: $\mu = 1.69$, $\sigma = 0.62$

Low-load series

The following section contains the complete raw data outputs for detection speed from each subject in the low-load series. For each graph, the presented symbol is indicated on the horizontal axis with detection time in seconds on the vertical. Absent data points indicate that the observer response failed to meet the criteria for accurate symbol recognition. Those criteria are:
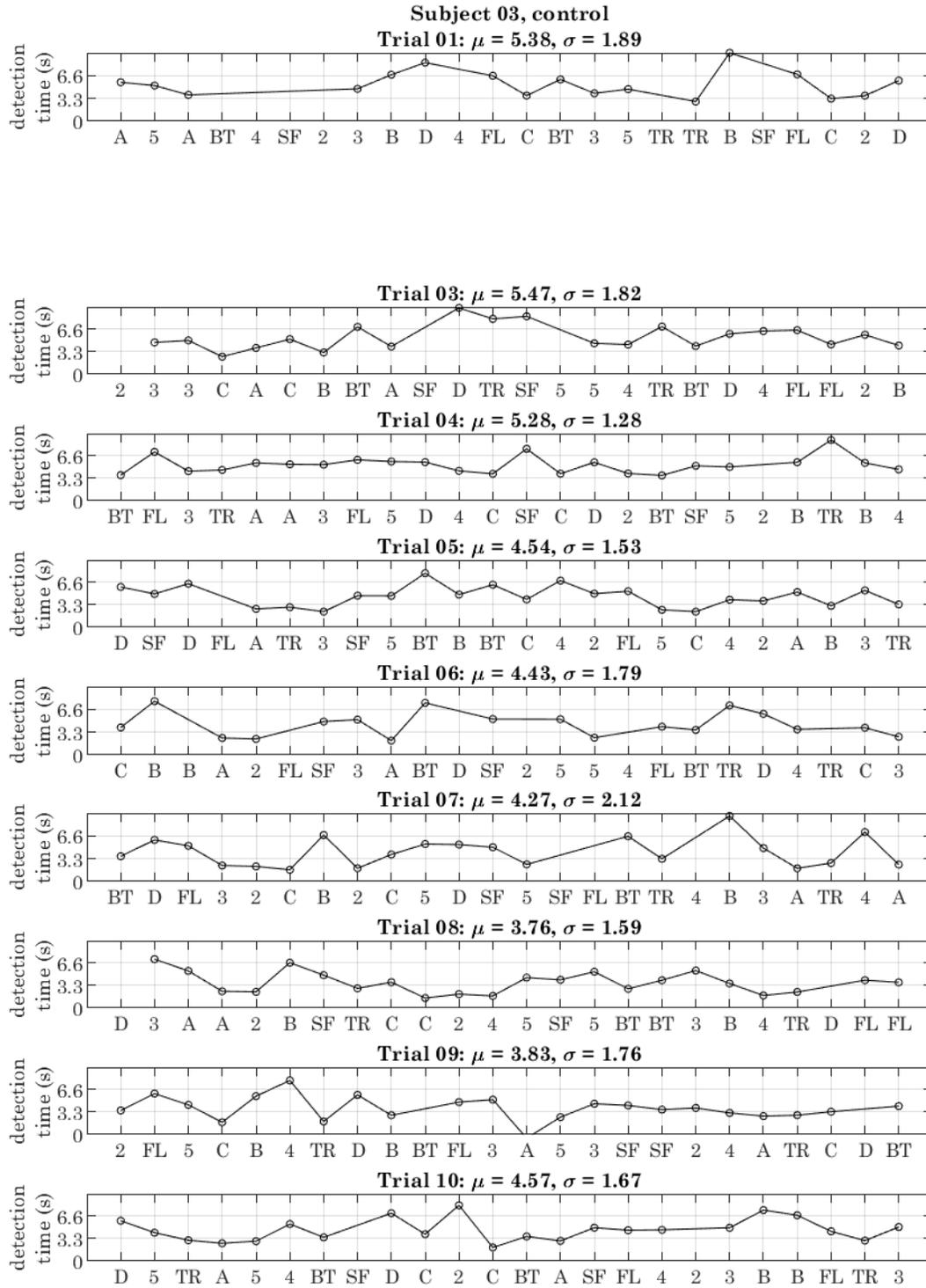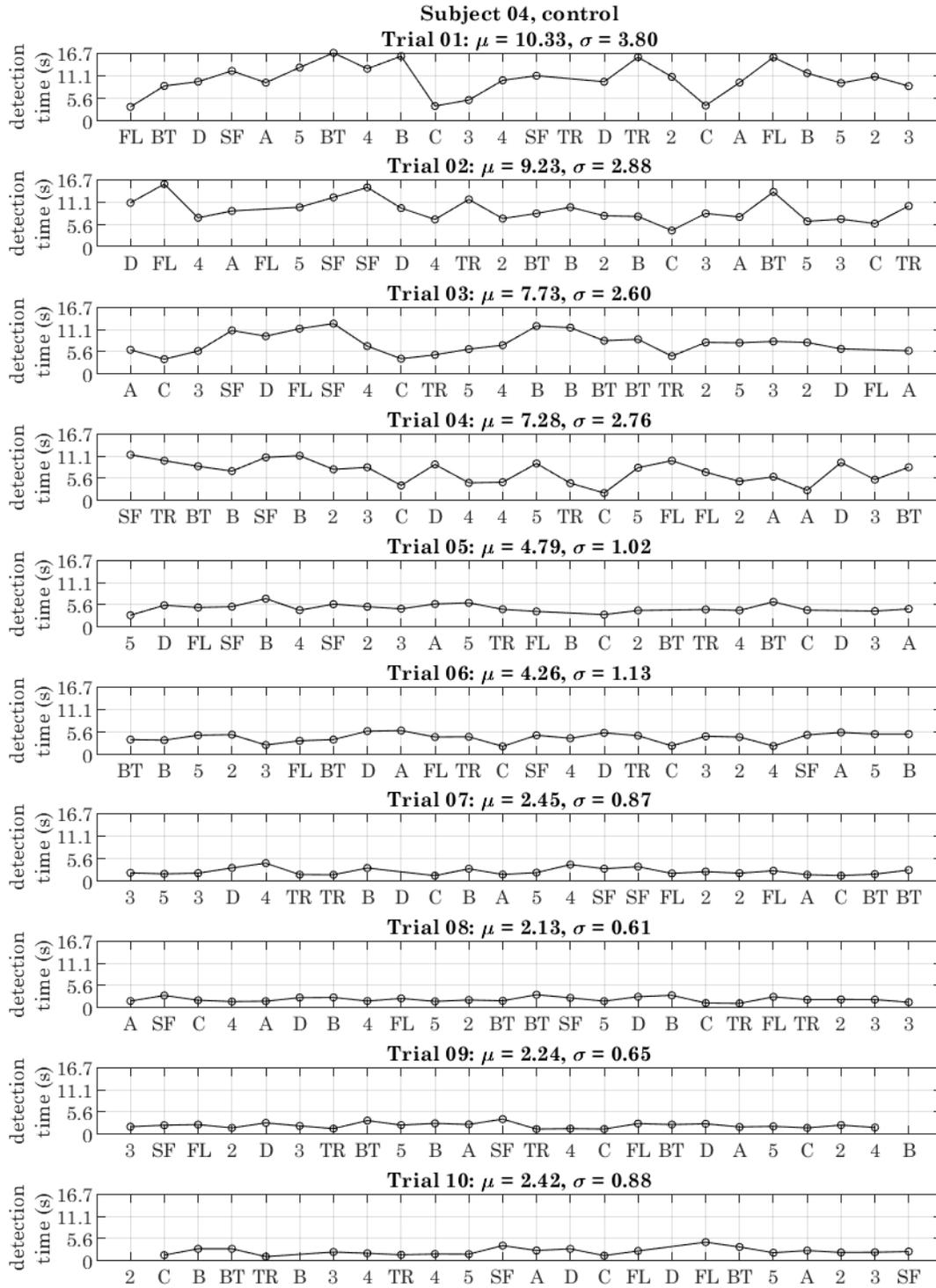
1. The codex block is accurately identified.
2. Detection was indicated by button press while the codex block was present.

This data is arranged longitudinally in a single graphic, along with data for the control and high-load series, for each observer (Subject 1 in Figure 32, Subject 2 in Figure 33) along with fitted trends. For full analysis of trends in detection speed over successive trials, please see section "7.4.1 Detection speeds continue to improve despite increasing scene complexity."

Subject 01, lowload

Trial 01: $\mu = 2.93$, $\sigma = 1.33$

SF 4 3 D B 5 BT C 5 4 SF FL D 2 TR 2 3 BT TR C A FL A B

Trial 02: $\mu = 2.21$, $\sigma = 0.89$

5 D A B 4 SF A C BT 3 BT TR SF TR 4 2 C 5 3 2 FL FL B D

Trial 03: $\mu = 2.32$, $\sigma = 1.01$

5 FL SF 4 B TR C SF 3 A 5 D FL B 4 3 D 2 C 2 TR BT BT A

Trial 04: $\mu = 2.32$, $\sigma = 1.65$

FL FL TR 4 2 4 C 3 A BT 5 SF B BT B 5 D D SF 2 A C 3 TR

Trial 05: $\mu = 2.42$, $\sigma = 1.23$

3 TR D FL 4 A BT B 5 3 TR D 2 5 SF 2 A B C C 4 FL BT SF

Trial 06: $\mu = 1.94$, $\sigma = 0.57$

3 5 2 D FL D C A C TR 5 B TR FL BT SF 3 A 4 SF 4 BT B 2

Trial 07: $\mu = 1.98$, $\sigma = 0.79$

5 2 SF D 3 SF 4 3 4 FL C D BT 2 FL 5 B B BT C A TR A TR

Trial 08: $\mu = 1.93$, $\sigma = 0.67$

4 5 2 SF 5 BT D 3 C 3 FL B C 2 FL B A BT A D TR SF 4 TR

Trial 09: $\mu = 0.68$, $\sigma = 0.20$

TR FL SF FL 3 C A 2 3 SF 2 4 4 B D BT BT A 5 B 5 TR D C

Trial 10: $\mu = 0.59$, $\sigma = 0.12$

B FL 2 5 FL 2 3 C D SF 4 SF C TR 5 3 B TR D BT BT A 4 A

Subject 02, lowload

Trial 01: $\mu = 2.18$, $\sigma = 0.62$

SF 4 3 D B 5 BT C 5 4 SF FL D 2 TR 2 3 BT TR C A FL A B

Trial 02: $\mu = 2.05$, $\sigma = 0.82$

5 D A B 4 SF A C BT 3 BT TR SF TR 4 2 C 5 3 2 FL FL B D

Trial 03: $\mu = 2.12$, $\sigma = 0.50$

5 FL SF 4 B TR C SF 3 A 5 D FL B 4 3 D 2 C 2 TR BT BT A

Trial 04: $\mu = 1.70$, $\sigma = 0.34$

FL FL TR 4 2 4 C 3 A BT 5 SF B BT B 5 D D SF 2 A C 3 TR

Trial 05: $\mu = 1.63$, $\sigma = 0.48$

3 TR D FL 4 A BT B 5 3 TR D 2 5 SF 2 A B C C 4 FL BT SF

Trial 06: $\mu = 1.63$, $\sigma = 0.48$

3 5 2 D FL D C A C TR 5 B TR FL BT SF 3 A 4 SF 4 BT B 2

Trial 07: $\mu = 1.44$, $\sigma = 0.37$

5 2 SF D 3 SF 4 3 4 FL C D BT 2 FL 5 B B BT C A TR A TR

Trial 08: $\mu = 1.42$, $\sigma = 0.46$

4 5 2 SF 5 BT D 3 C 3 FL B C 2 FL B A BT A D TR SF 4 TR

Trial 09: $\mu = 1.52$, $\sigma = 0.40$

TR FL SF FL 3 C A 2 3 SF 2 4 4 B D BT BT A 5 B 5 TR D C

Trial 10: $\mu = 1.59$, $\sigma = 0.48$

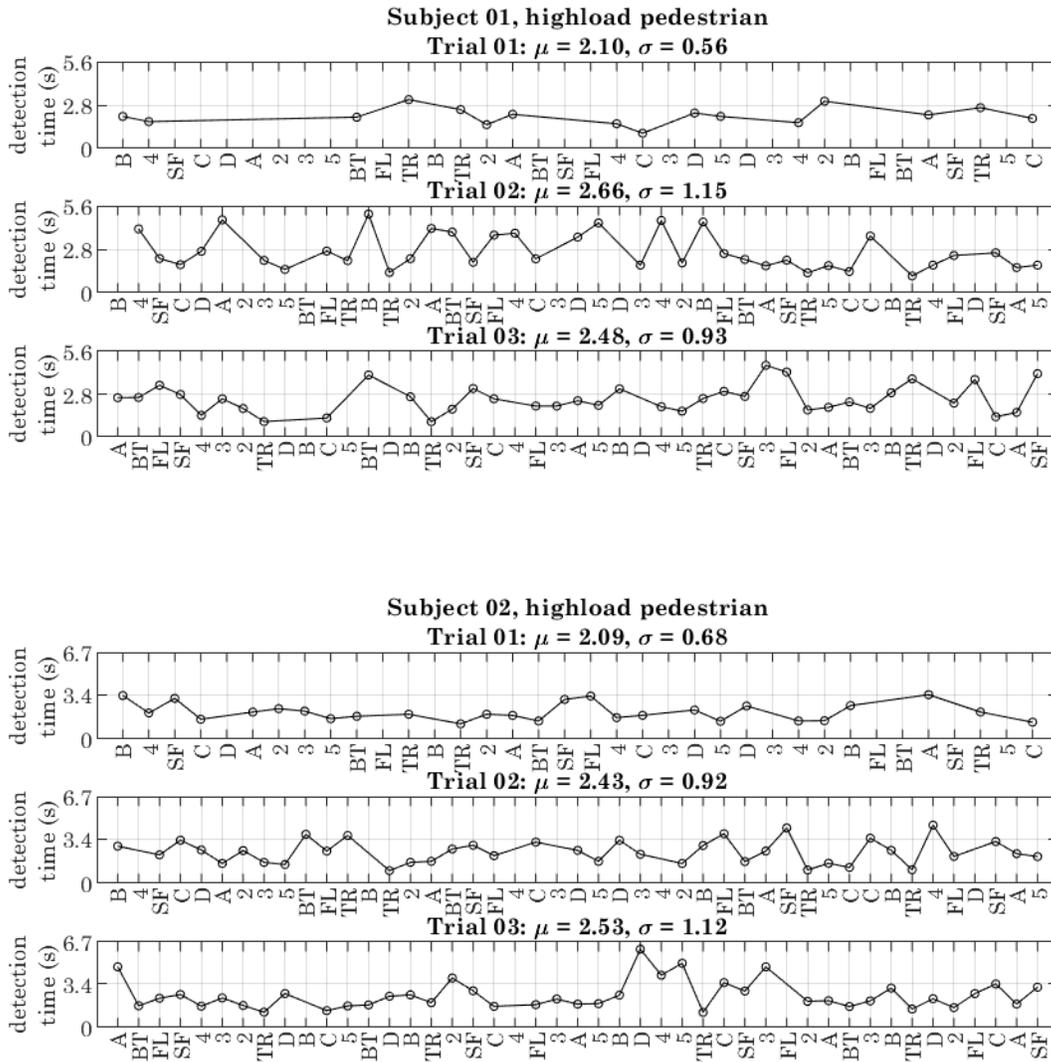B FL 2 5 FL 2 3 C D SF 4 SF C TR 5 3 B TR D BT BT A 4 A

High-load series: pedestrian

The following section contains the complete raw data outputs for detection speed from each subject in the high-load series, pedestrian POV footage. For each graph, the presented symbol is indicated on the horizontal axis with detection time in seconds on the vertical. Absent data points indicate that the observer response failed to meet the criteria for accurate symbol recognition. Those criteria are:

1. The codex block is accurately identified.
2. Detection was indicated by button press while the codex block was present.

This data is arranged longitudinally in a single graphic, along with data for the control and high-load series, for each observer (Subject 1 in Figure 32, Subject 2 in Figure 33) along with fitted trends. For full analysis of trends in detection speed over successive trials, please see section "7.4.1 Detection speeds continue to improve despite increasing scene complexity."

Subject 01, highload pedestrian

Trial 01: $\mu = 2.10$, $\sigma = 0.56$

Trial 02: $\mu = 2.66$, $\sigma = 1.15$

Trial 03: $\mu = 2.48$, $\sigma = 0.93$

Subject 02, highload pedestrian

Trial 01: $\mu = 2.09$, $\sigma = 0.68$

Trial 02: $\mu = 2.43$, $\sigma = 0.92$
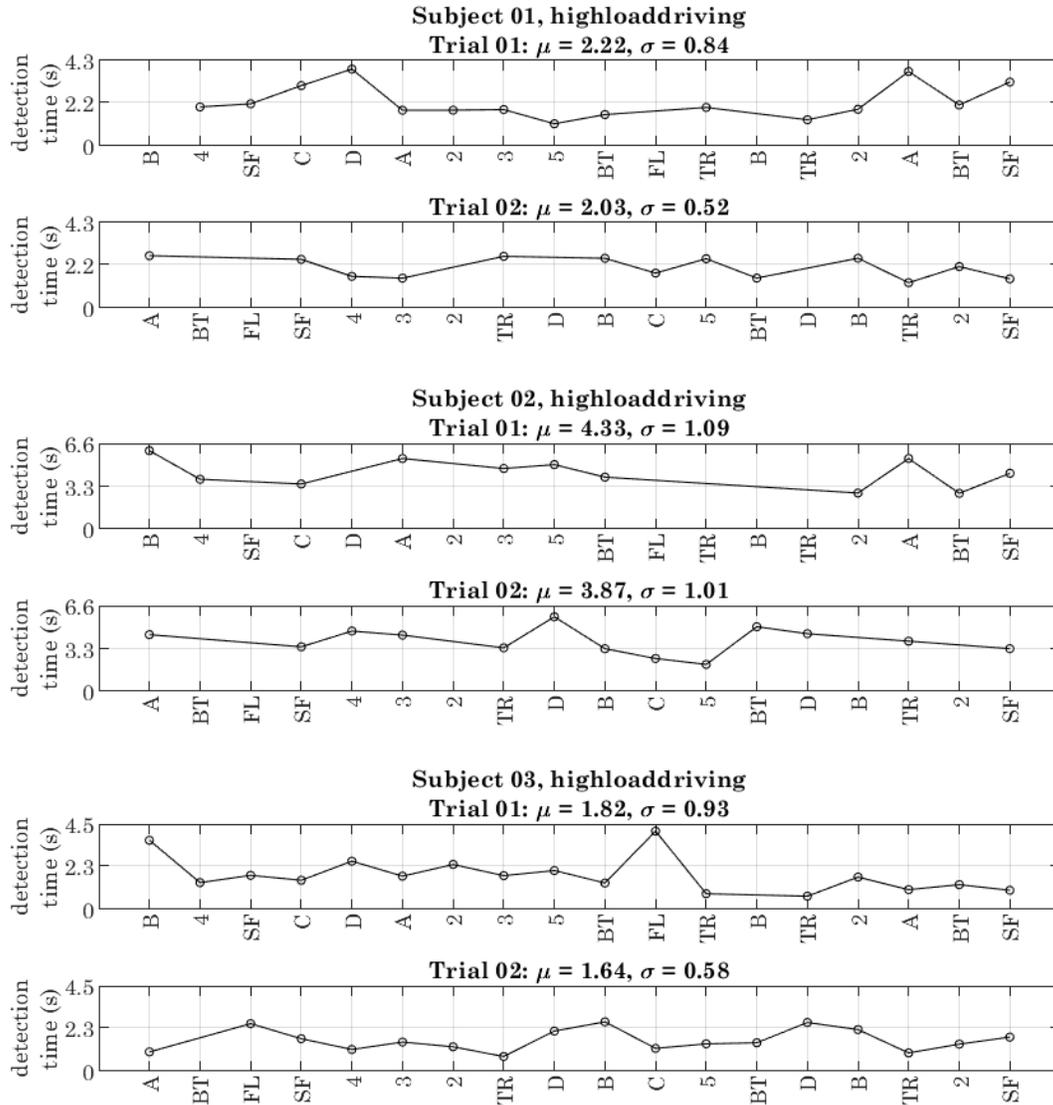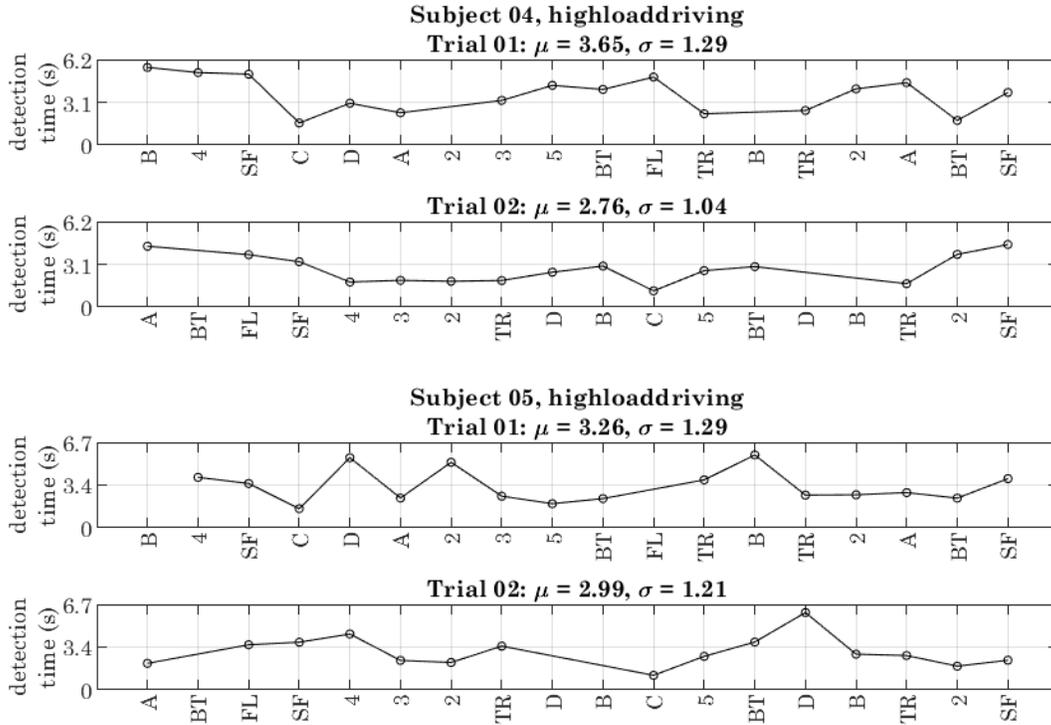
Trial 03: $\mu = 2.53$, $\sigma = 1.12$

High-load series: driving

The following section contains the complete raw data outputs for detection speed from each subject in the high-load series, automotive cockpit POV. For each graph, the presented symbol is indicated on the horizontal axis with detection time in seconds on the vertical. Absent data points indicate that the observer response failed to meet the criteria for accurate symbol recognition. Those criteria are:

1. The codex block is accurately identified.
2. Detection was indicated by button press while the codex block was present.

This data is arranged longitudinally in a single graphic for subjects 1 and 2 (Figure 32 and Figure 33) and subjects 3, 4, and 5 (Figure 34). For full analysis of trends in detection speed over successive trials, please see section "7.4.1 Detection speeds continue to improve despite increasing scene complexity" and "7.4.2 Visual learning outcomes from control series persist in complex environments both with and without additional exposure."

Subject 01, highloaddriving
Trial 01: μ = 2.22, σ = 0.84

Trial 02: μ = 2.03, σ = 0.52

Subject 02, highloaddriving
Trial 01: μ = 4.33, σ = 1.09

Trial 02: μ = 3.87, σ = 1.01

Subject 03, highloaddriving
Trial 01: μ = 1.82, σ = 0.93

Trial 02: μ = 1.64, σ = 0.58

Subject 04, highloaddriving
Trial 01: $\mu = 3.65$, $\sigma = 1.29$

Trial 02: $\mu = 2.76$, $\sigma = 1.04$

Subject 05, highloaddriving
Trial 01: $\mu = 3.26$, $\sigma = 1.29$

Trial 02: $\mu = 2.99$, $\sigma = 1.21$

Fitted trends: low-load, high-load

Exponential and linear trends for the low- and high-load series data (respectively) are summarized in the following figures. Raw data is indicated in blue while the fitted trend is shown in red. These fitted trends are also summarized and combined into a single, longitudinal graphic for each corresponding subject in Figure 32, Figure 33, and Figure 34. In each graphic below, the coefficients of the fitted trend (exponential decay or linear) are indicated for each subject, as well as mean (μ) and standard deviation (σ).

Symbol detection rates for subjects {1,2} over 10 successive trials in the low-load series (pedestrian POV environment).
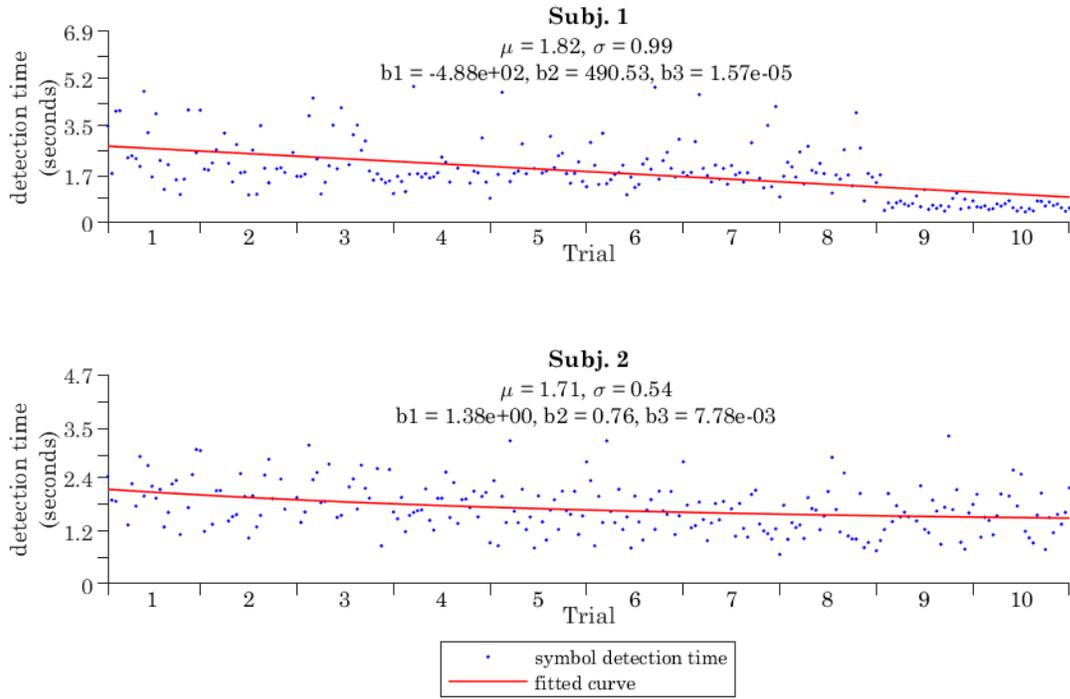
Symbol detection rates for subjects {1,2} over 3 successive trials in the high-load series (pedestrian POV environment).
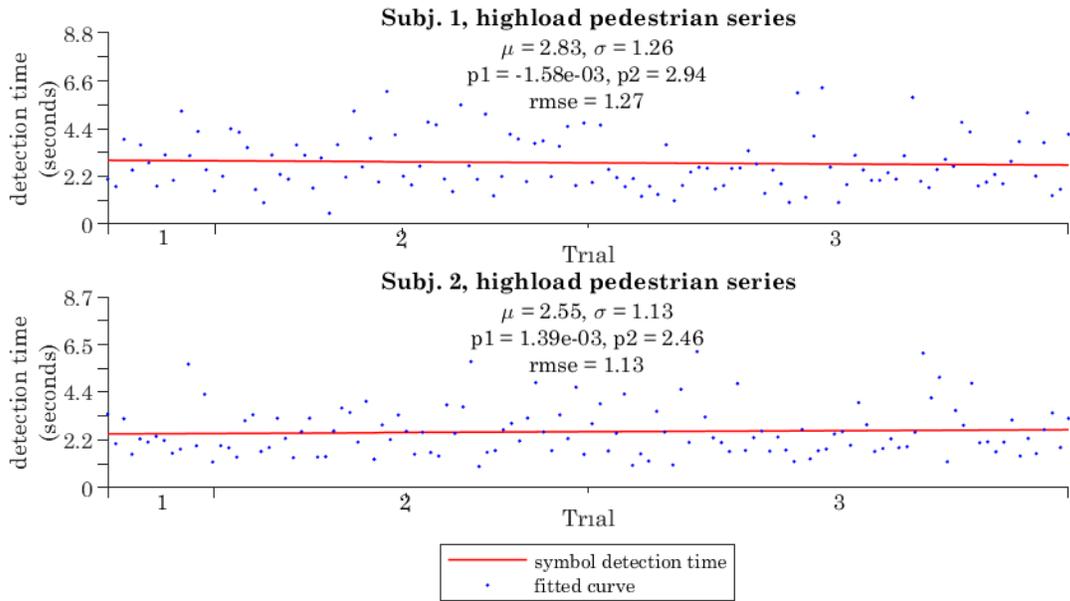
Symbol detection rates for subjects {1,2,3,4,5} over 2 successive trials in the high-load series (automotive cockpit environment)

Appendix B: Supplemental results