POSITION REPRESENTATIONS OF MOVING OBJECTS ALIGN WITH REAL-TIME POSITION IN THE EARLY VISUAL RESPONSE

Philippa A. Johnson*
University of Melbourne
pajohnson@student.unimelb.edu.au

Tessel Blom
University of Melbourne
tesselblom@gmail.com

Simon van Gaal
University of Amsterdam
simonvangaal@gmail.com

Daniel Feuerriegel
University of Melbourne
dfeuerriegel@unimelb.edu.au

Stefan Bode
University of Melbourne
sbode@unimelb.edu.au

Hinze Hogendoorn
University of Melbourne
hhogendoorn@unimelb.edu.au

June 2022

ABSTRACT

When localising a moving object, the brain receives outdated sensory information about its position, due to the time required for neural transmission and processing. The brain may overcome these fundamental delays through predictively encoding the position of moving objects using information from their past trajectories. In the present study, we evaluated this proposition using multivariate analysis of high temporal resolution electroencephalographic data. We tracked neural position representations of moving objects at different stages of visual processing, relative to the real-time position of the object. During early stimulus-evoked activity, position representations of moving objects were activated substantially earlier than the equivalent activity evoked by unpredictable flashes, aligning the earliest representations of moving stimuli with their real-time positions. These findings indicate that the predictability of straight trajectories enables full compensation for the neural delays accumulated early in stimulus processing, but that delays still accumulate across later stages of cortical processing.

Keywords motion extrapolation · prediction · latency · neural delays · EEG

1 Introduction

Localising objects within a dynamic world is a primary function of the visual system: catching prey, escaping predators, and avoiding moving objects (e.g., falling rocks) are all crucial to survival. This task is complicated by delays that accumulate during the neural transmission of visual information from the eyes to the brain. As a result, the visual cortex only has access to outdated sensory information. Furthermore, additional delays accumulate during subsequent cortical processing of visual information. The world will inevitably change during this time, so how can the brain overcome this fundamental problem and keep up with an ever-changing world?

Several lines of evidence suggest that the brain can compensate for neural transmission delays through prediction: using information from the past to predict what is happening in the present[1]. Indeed, many researchers consider prediction to be a core objective of the central nervous system[2;3]. In the field of motion perception, there is evidence that predictive processes help to compensate both for the neural delays incurred before visual input reaches the visual cortex and for the delays incurred during subsequent cortical processing[4;5;6;7].

For example, neurophysiological recordings in animals reveal motion extrapolation mechanisms as early as the retina[4;8;9;10;11]. By responding to the leading edge of moving stimuli, retinal ganglion cells extrapolate the represented position of those stimuli, and are thought to transmit these extrapolated representations to visual cortex, thereby compensating for some of the lag that arises during transmission. These pre-cortical extrapolation mechanisms should effectively allow primary visual cortex to represent the position of a moving object with reduced latency, as observed in both cat and macaque V1[5;6]. The existence of these extrapolation mechanisms opens the possibility that transmission delays on the way to visual cortex might be partially or fully compensated, allowing the early visual system to represent moving objects on predictable trajectories closer to their real-time locations.

It is unclear whether similar mechanisms operate along the cortical visual processing hierarchy to compensate for additional delays that accumulate as visual information is processed. On the one hand, there is suggestive evidence that position representations in areas V4[7] and V5[12] are shifted for moving objects, potentially reflecting the effect of motion extrapolation in those areas. That interpretation is consistent with recent fMRI[13;14], theoretical[15] and psychophysical[16] work suggesting that motion extrapolation mechanisms operate at multiple levels of the visual system. On the other hand, shifted position representations in higher areas might simply result from those areas inheriting extrapolated information from upstream areas such as V1. To our knowledge, no study to-date has investigated how the represented position of a smoothly moving object evolves over time as visual information about that object flows along the visual hierarchy.

Here, we address this question by using time-resolved EEG decoding to probe the position representations of smoothly moving objects across all levels of the human visual system in real-time. We show that early position representations of moving objects are in close alignment with the veridical position of the object, providing the first direct evidence in humans that extrapolation processes allow the early visual system to localise moving objects in real-time. We further show that during the course of cortical visual processing, position representations increasingly lag behind real-time stimulus position as information progresses through the visual hierarchy. This suggests that delay compensation is primarily achieved at very early stages of stimulus processing, and that subsequent cortical visual areas do not implement further compensation for neural delays. Nevertheless, this early compensation ensures that the represented position of a moving object throughout the entire visual hierarchy is far more up-to-date than could be expected on the basis of the latencies of neural responses to static objects. These findings demonstrate the existence of significant predictive processing during motion perception, but constrain any predictive mechanisms to acting relatively early in processing.

2 Results

Twelve observers viewed sequences of black discs that were either flashed in one of 37 possible positions on a hexagonal grid (static trials), or smoothly moved through a series of positions on the grid along a straight trajectory (motion

trials, Figure 1). Static trials were presented 252 times per position, and each of the 42 motion vectors was presented 108 times. EEG data were recorded over six testing sessions and analysed offline (see Methods). Multivariate pattern classifiers (linear discriminant analysis) were trained to discriminate stimulus position for all pairwise combinations of positions, using EEG activity evoked by static stimuli in those positions. These classifiers were subsequently tested on EEG data recorded during an independent subset of static trials, or during motion trials. Results of this classification analysis were combined to estimate the likelihood of the stimulus being present in each of the possible stimulus positions, $p_s(i)$ for $i \in \{1, 2, \ldots 37\}$, where s is the presented position. From this, we traced the evolution over time of the estimated likelihood of the stimulus being present in the position in which it was actually presented (static trials) or moved through (motion trials), $p_s(s)$, hereafter referred to as the *stimulus-position likelihood*.

This analysis was repeated for multiple combinations of training timepoint (i.e. time after onset of a static stimulus) and test timepoint (i.e. time after presentation of a static stimulus or within a motion vector). Using different training timepoints allowed us to probe neural representations at different levels of the visual hierarchy, and testing at multiple timepoints allowed us to characterise how information flows through those levels over time during the epoch of interest[17]. In this way, we were able to evaluate whether the neural position representation of a moving object flows through the visual hierarchy at the same latency as the position representation of a static flash. Additionally, this allowed us to evaluate how much the position representation of the moving object lags behind that object's physical position.

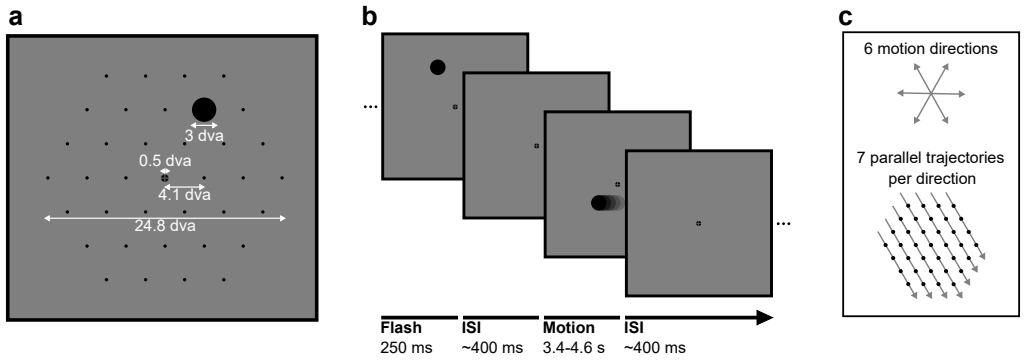

Figure 1: ***Stimuli in static and motion trials. a)*** *Stimulus configuration. Stimuli were presented in a hexagonal grid. In static trials, a black circle was shown centred in one of the 37 positions (marked by black dots, not visible during the experiment). In motion trials, the same stimulus moved at 10.36 degrees visual angle/second in a straight line through the grid. A fixation point was presented in the centre of the screen and the background was 50% grey. All measurements are in degrees visual angle (dva).* ***b)*** *Trial structure. A trial consisted of a black circle flashed in one position for 250ms (static trials) or moving in a straight line for between 3350 and 4550ms (motion trials). Trials were randomly shuffled and presented separated by an inter-stimulus interval randomly selected from a uniform distribution between 350ms and 450ms.* ***c)*** *Motion trials. The moving stimulus travelled along one of 42 possible straight trajectories through the grid: six possible stimulus directions along the hexagonal grid axes with seven parallel trajectories for each direction. The moving stimulus passed through four to seven flash locations, depending on the eccentricity of the trajectory.*

2.1 Decoding positions of static stimuli

First, we investigated the ability of classifiers to discriminate the presented position of static flashes based on the EEG signal. Figure 2a shows average classification accuracy across pairwise combinations of positions over time, grouped by distance between the two positions. Classifiers were trained and tested using data from the same timepoints. As expected, the performance of pairwise classifiers improved with increasing stimulus separation. This is due to the retinotopic organisation of visual cortex; stimuli elicit more distinct patterns of activity when they are further apart.

Pairwise classification results were combined to calculate the stimulus-position likelihood. We then averaged across all stimulus positions and participants. This likelihood was compared to a permuted null-distribution to establish whether

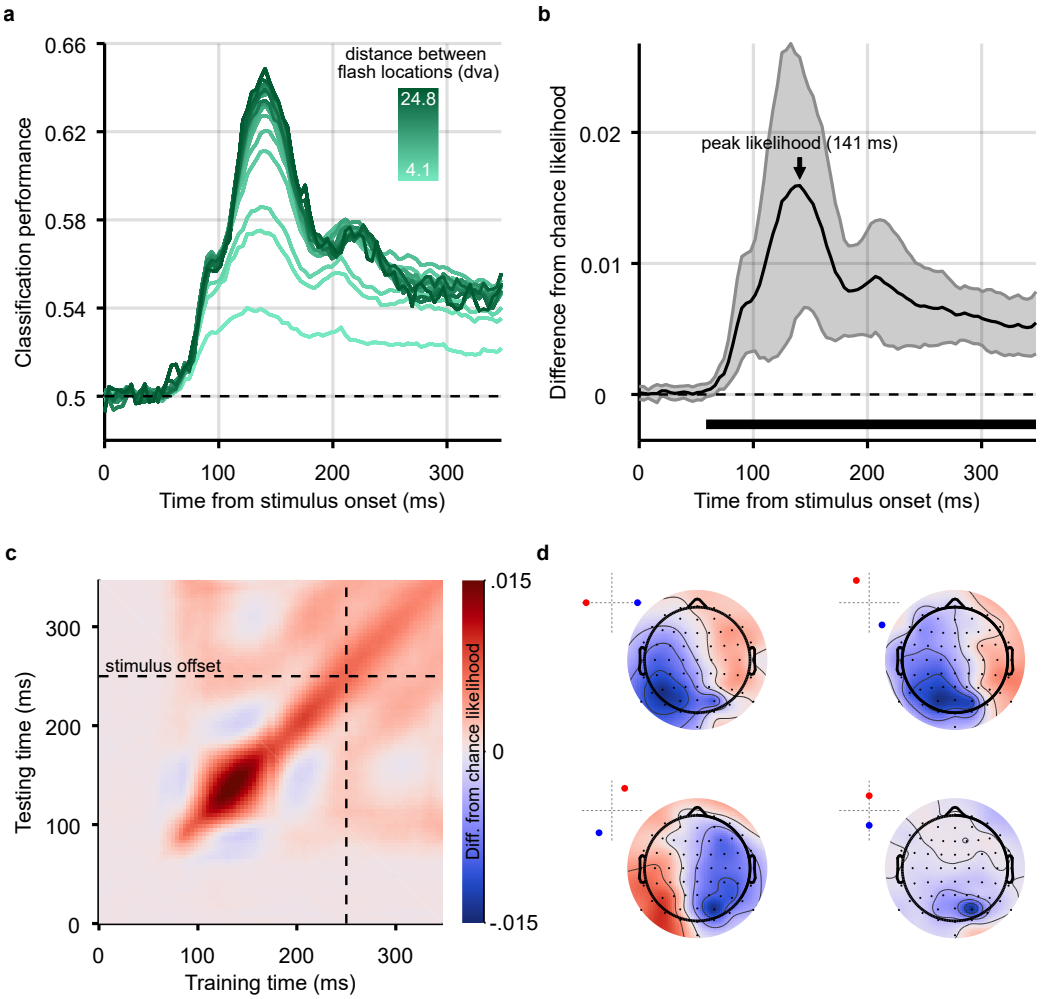

Figure 2: ***Classification results for decoding the position of static stimuli.*** ***a)*** *Group-level pairwise classification performance of static stimulus position discrimination sorted by distance between stimulus positions (separate lines). Classifiers were trained and tested on matched timepoints from 0-350ms (i.e. the time diagonal).* ***b)*** *Timepoints along the time diagonal at which likelihood of the stimulus being in the presented position (stimulus-position likelihood) is significantly above chance ($p<.05$, cluster-based correction applied) are marked by the bar above the x-axis. The stimulus-position likelihood was significantly above chance from 58ms onward. Shaded error bars show one standard deviation around the mean across observers. Chance level has been subtracted from all likelihoods to demonstrate the divergence from chance, in this graph and all others showing stimulus-position likelihood.* ***c)*** *Stimulus-position likelihood was calculated from classification results at each combination of training and test times. Results averaged across all stimulus positions and participants are displayed as a temporal generalisation matrix (TGM).* ***d)*** *Topographic maps show participant-averaged topographic activity patterns used by classifiers to distinguish stimulus positions at 141ms post stimulus onset, the time of peak decoding (marked by an arrow on panel B). Insets in the top left of each scalpmap show which two stimulus positions the classifier has been trained to discriminate. Scalp maps were obtained by combining classification weights with the relevant covariance matrix. As expected, for all four comparisons, activation was predominantly occipital and, when the stimulus positions were on either side of the vertical meridian, lateralised.*

it was significantly above chance at each timepoint (Figure 2b, see Methods). The stimulus-position likelihood was above chance starting at 58ms after stimulus onset.

To assess whether position-related information was stable or variable across the time-course of the visual evoked response, the classification analysis was generalised across time[17]: classifiers were trained and tested at all combinations of timepoints. Figure 2c shows the resulting temporal generalisation matrix (TGM), averaged across all stimulus po-

sitions and participants. The TGM was typical of position decoding plots seen in previous work [18]. Finally, Figure 2d shows topographic maps of activation which contributed to classification of stimulus position [19]; these show that the relevant signal was mainly recorded from occipital electrodes, suggesting a source within the visual cortex, as expected.

2.2 Decoding positions of moving stimuli

To decode the position of moving stimuli, we again trained classifiers on pairwise combinations of static stimuli, then applied these classifiers to EEG data recorded during motion trials. An illustration of each step in the analysis of motion trials is shown in Figure 3. As before, the stimulus-position likelihood was calculated, this time at each timepoint during each motion epoch.

We considered EEG epochs from 500ms before to 500ms after the timepoint at which the moving stimulus was exactly in each possible static stimulus location. This time-window was chosen to be broad enough to capture stimulus evoked activity as the stimulus approached and receded from each position (moving from one position to the next took 400ms). We then averaged the time-course of stimulus-position likelihoods across all six motion directions and 37 stimulus positions. The first position along each trajectory was excluded due to observed strong EEG responses to the initial onset of the stimulus.

The TGM derived from classifiers trained on static trials and tested on motion trials (Figure 3 step 3) revealed that classifiers trained on timepoints from around 100ms were able to decode the position of moving objects. To identify timepoints at which classification was significantly above chance, we considered the performance of classifiers trained and tested on matching timepoints (diagonal of the TGM). Permutation testing revealed that decoding was significantly above chance for timepoints between 102 and 180ms (Figure 3 step 4). Note that because we are investigating possible latency differences between the neural response to static and moving stimuli, maximal decoding is likely achieved off-diagonal, making this a conservative analysis choice.

Although the average stimulus-position likelihood was smaller in magnitude for moving stimuli compared to static stimuli, we observed that the location-specific neural response to motion over time was characterised by a gradual increase of the likelihood of the stimulus being present as the stimulus approached the centre of the position, then a decrease as the stimulus moved away on the other side. This is illustrated in Figure 3 step 5, and is similar to the pattern of activity found in response to a moving bar with direct recordings from cat V1 [5].

2.3 Latency of position representations of moving stimuli

To investigate the latency at which neural position representations are activated for moving objects, we calculated the timepoint at which the peak stimulus-position likelihood was reached during motion sequences. Again, this was repeated for different training times as a proxy for different stages of neural processing. The time to peak likelihood in the test data for each training timepoint reflects the time at which the location-specific activity is most similar in the train and test set, assumed to be the time that the brain is representing the moving stimulus at the centre of a flash position. We use *peak* likelihood, as opposed to onset or a peak percentage, as the latency measure due to possible variations in receptive field (RF) size over the course of visual processing. As time elapses during stimulus processing, visual information reaches visual areas further up the processing hierarchy, which contain stimulus-selective neurons with larger RFs [20;21]. This would mean that a moving stimulus would enter the RF earlier in these later-activated brain regions. Looking at the peak neural response avoids this problem, because peak response would be expected when the stimulus is at the centre of the RF, irrespective of RF size.

To establish the latency with which the position of a moving object is represented at different stages of visual processing, we identified the timepoint at which our classification analysis yielded maximum stimulus-position likelihood. To overcome the noise of individual data points, we fit a Gaussian to the observed time-course of the calculated likelihood

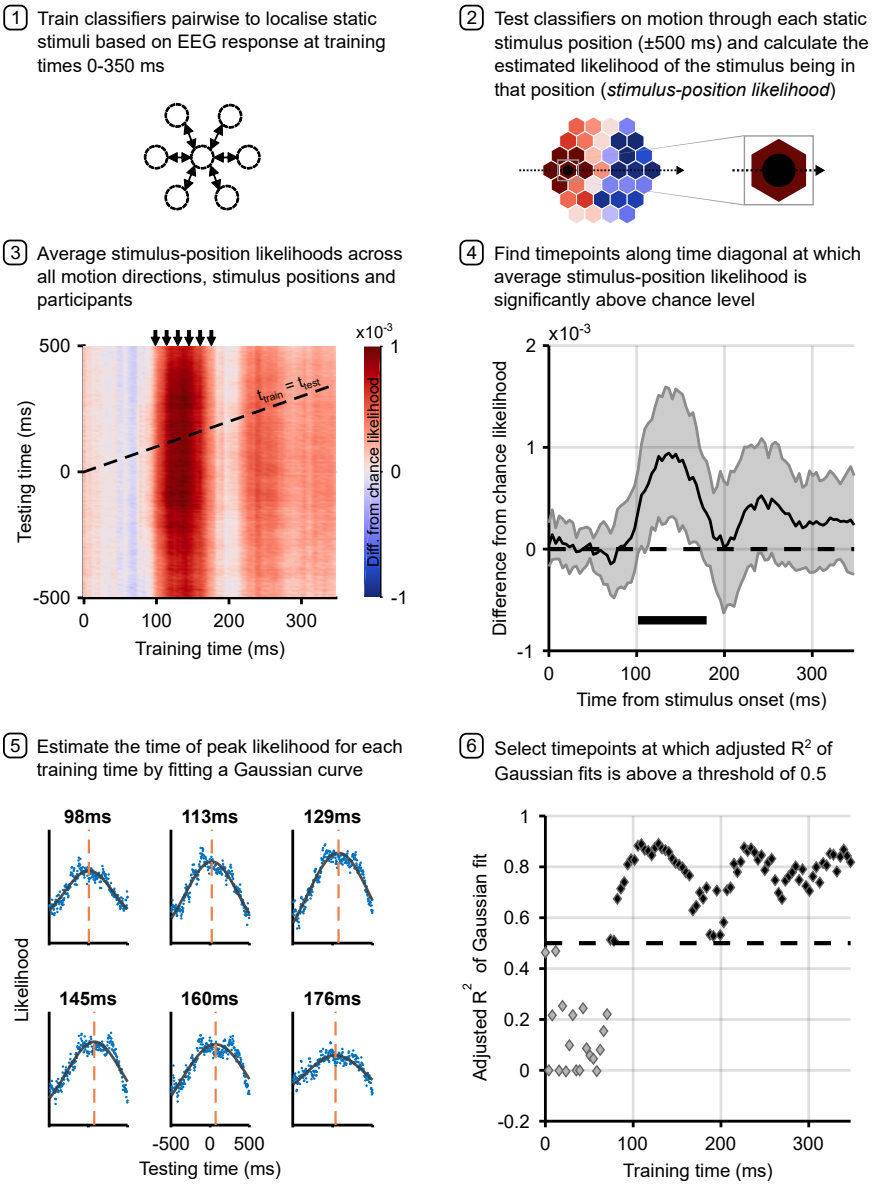

Figure 3: ***Analysis pipeline for motion trials.*** *Panels describe steps in calculating the time to peak stimulus-position likelihood in motion trials, including graphs of relevant data for each step. Steps 1 and 2 describe the classification analysis applied to obtain the stimulus-position likelihoods. The figure in step 3 shows the group-level temporal generalisation matrix for training on static stimuli and testing on moving stimuli. The black dotted line shows the 'diagonal' timepoints, where the time elapsed since the moving stimulus was at the flash position equals the training time. Step 4 shows timepoints along this diagonal at which the stimulus-position likelihood was significantly above chance, as established through permutation testing. Significance is marked by the solid black line above the x-axis; the likelihood is significantly above chance from 102-180ms. Shaded error bars show one standard deviation around the mean. The figure in step 5 shows the same data as step 3 for selected training times (arrows above TGM correspond to subplot titles). Each subplot shows a vertical slice of the TGM. Blue points show data, to which we fit a Gaussian curve (black lines) to estimate the time of peak likelihood for each training time (dashed orange lines). These are the data points plotted in Figure 4a and b. Step 6 shows adjusted R^2 of Gaussian fits for each training timepoint. A cutoff of 0.5 was used to select timepoints at which the Gaussian fit meaningfully explains the pattern of data.*

averaged across participants, separately for each training time. There were four free parameters:

$$p_s(s) = b_1 \exp\left(-\frac{t_{test} - b_2^2}{2b_3^2}\right) + b_4.$$

The parameter of interest is b_2, which describes the horizontal shift of the peak of the Gaussian. Adjusted R^2 of these fits can be found in Figure 3 step 6. For training timepoints later than ∼80ms, the Gaussian curves provided a very good fit to the evolution of stimulus-position likelihood over time, with R^2 values over 0.5. Although the window of significant cross-classification of static stimuli to moving stimuli is restricted (Figure 3, step 4), the sustained high adjusted R^2 values indicate that even for training times at which the stimulus-position likelihood was close to chance level, the likelihood increased and decreased as the stimulus traversed each flash location.

Figure 4a shows the time to peak likelihood for motion across all training timepoints at which adjusted R^2 exceeded a minimum value of 0.5. The choice of R^2 cutoff is essentially arbitrary, but the pattern of points in Figure 3 step 6 shows that this selection is relatively robust to changes in the cutoff value. Up to ∼150ms training time, the time to peak likelihood increases with increasing training time. This follows the same pattern as the static trials (see Figure 4b), where earlier representations of the stimulus (i.e. early training times) were activated at a shorter latency in the testing epoch than later representations of the stimulus. This sequential pattern is consistent with the first feed-forward sweep of stimulus-driven activation. As information flows through the visual processing hierarchy, representations of the stimulus will gradually change over time. The order of these changes appeared to be consistent between static and motion trials.

This pattern subsequently reverses between 150 and 200ms, indicating that hierarchically later representations were activated at a shorter latency. Finally, from a training time of ∼250ms the time to peak likelihood was stable at approximately 50ms. The non-monotonic relationship between training time and time to peak likelihood could emerge because there was variable compensation for neural delays at different training times. Perhaps more likely, this pattern could reflect feedforward and feedback sweeps of activity in the visual cortex: the feedforward sweep activates sequential representations, and information flowing backward along the hierarchy, reactivates the same activity patterns in reverse order[22]. The timescale of this wave of activity was in line with previous findings from Dijkstra et al.[22], who showed approximately 10Hz oscillations evoked by face/house stimuli. Additionally, previous TMS, MEG and fMRI results suggest that 150ms is a reasonable estimate for the time it takes for visual information to reach later visual processing areas, such as V5/MT+[23;24;25;26;27;28]. If this later activity (>150ms training time) does indeed reflect feedback processing, then, for these later timepoints, the latency measure we have calculated might not be informative about the time necessary to first represent the moving object, because the initial activations and the reactivations are indistinguishable.

In order to confirm which timepoints predominantly correspond to feedforward processes rather than feedback, we applied piecewise regression, as implemented in the Shape Language Modeling toolbox[29]. In this approach, several polynomials are smoothly joined together at 'knots'. Placement of knots, at the start and end of each segment, is optimised by reducing root mean squared error. We fit straight lines, and varied the number of knots between four and seven, in order to identify the optimal number. The best piecewise regression fit, with six knots, is shown in Figure 4a. Corroborating our observation, the first internal knot was placed at 158ms. We took this inflection point as the end of the initial feedforward sweep of information through the visual cortex. The piecewise regression revealed further positive and negative slopes, suggesting that feedforward and feedback sweeps of activity continue during later stimulus processing[22].

2.4 Latency advantages during feedforward processing

Having identified the timepoints during the motion epochs corresponding to early feedforward processing, we further investigated the relationship between training time and the time of peak stimulus-position likelihood. Figure 4b shows the peak time for moving objects during the feed-forward sweep, along with the first segment of the fitted piecewise linear regression. We were interested in comparing the latency and time-course of stimulus-related processing of static

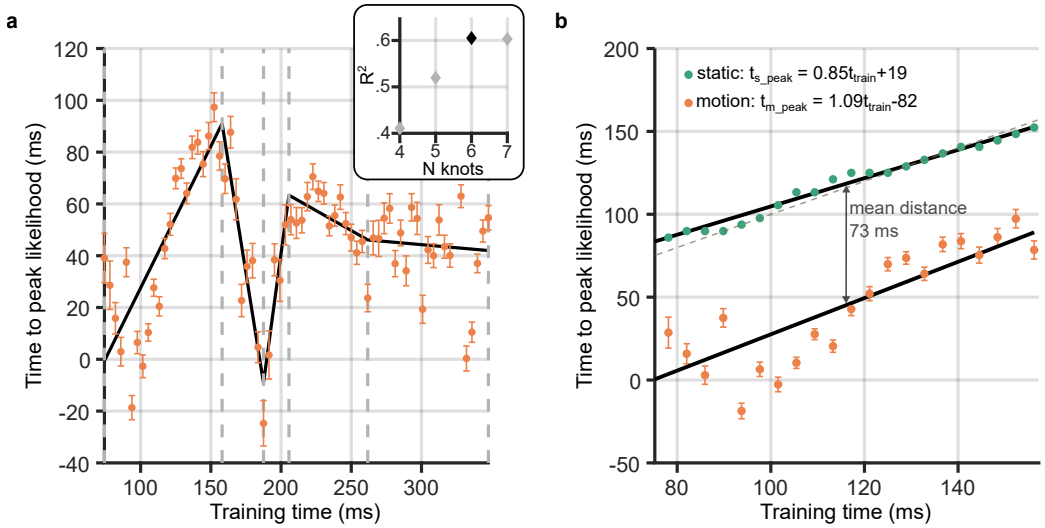

Figure 4: ***Neural response latency during motion processing.*** ***a)*** *Latencies of the peak stimulus-position likelihood values during motion processing. The timepoint at which peak likelihood was reached is plotted against training time. Errorbars around points show bootstrapped 95% confidence intervals of the peak shift parameter of the Gaussian fit (see Figure 3 step 5). It can be observed that the peak time increases and decreases, then levels out. Points of inflection within this timeseries were identified using piecewise regression (shown in black). The number of inflection points, or knots, was established by comparing the R^2 of piecewise regression fits, as shown in the inset graph. It was determined that six knots was optimal; positions of these knots are marked by grey dotted lines on the main graph.* ***b)*** *Time to peak likelihood during the initial feedforward sweep of activity through the visual cortex. Displayed is a subset of points from those shown in panel A, corresponding to a restricted time-window between the first two knots, during which the first feedforward sweep of activity was most likely occurring. The dotted diagonal shows the 45° line, where the time of peak likelihood would equal the training time. Data points from static trials (green) should theoretically lie along this line, as, in this case, the training and test data were subsets of the same trials. Straight lines were fit separately for static and motion trials. Both lines had similar gradients, close to unity, indicating equivalent cumulative processing delays for static and motion trials within this training time-window. However, the intercept for motion was much earlier at -80ms. The mean distance between the two lines is marked, indicating that position representations were activated ∼70ms earlier in response to a moving stimulus compared to a flashed one in the same location. Time to peak likelihood at the beginning of the feedforward sweep was approximately 0ms, indicating near-perfect temporal alignment with the physical position of the stimulus.*

and motion trials in this restricted time-window, to establish whether the position of moving objects was predictively encoded. We therefore established time to peak likelihood for these same training timepoints in the static trials. Because the participant-averaged time-course of the stimulus-position likelihood for each training time was much less noisy for static than motion trials, the time of peak likelihood was computed as a simple maximum for each training time. Qualitatively, it can be observed that this lay along the diagonal of the TGM (Figure 2b, green points).

For the static stimuli, a linear fit relating training time and time to peak likelihood was very similar to the 45° line (Figure 4b, upper line), revealing that each representation of the stimulus was active at roughly the same time in the training and test data (Figure 4b, grey dashes). A linear fit to the static datapoints was significantly better than a constant model ($F(21,19) = 805.45$, $p = 1.24\text{x}10^{-17}$). The line which best described the relationship between static time to peak likelihood and training time was:

$$t_{s\_peak} = 0.85t_{train} + 19.$$

Both the intercept and gradient parameters were significantly different from zero ($t = 5.48$, $p = 2.32\text{x}10^{-5}$; $t = 28.38$, $p = 1.24\text{x}10^{-17}$). The line had a gradient close to one (95% CIs: 0.79 to 0.92) and a small intercept (95% CIs: 12 to 27ms), indicating only a small shift in the peak time between training and testing. This line fit indicates that the patterns of

activity on which the classifiers were trained were most similar to activity at approximately the same timepoint in the test trials. This is as expected, as train and test data are subsets of the same static trials.

In contrast, we found evidence of a shift in the latency of activation of representations of the moving stimulus. A regression line was also found for time to peak likelihood in the motion epochs ($F(21,19) = 40.91$, $p = 3.07\text{x}10^{-6}$), with equation:

$$t_{m\_peak} = 1.09t_{train} - 82.$$

Both parameters were again significantly different from zero (intercept: $t = -4.06$, $p = 6.12\text{x}10^{-4}$; gradient: $t = 6.40$, $p = 3.07\text{x}10^{-6}$). The 95% confidence intervals of the gradient overlapped with those of the flash line (flash: 0.79 to 0.92; motion: 0.73 to 1.45). This indicates that once position information was available in the cortex, successive cortical representations were sequentially activated along the same time-course for moving and static stimuli. In other words, delays that accumulate during cortical processing did not appear to be compensated when processing motion.

Importantly, however, the linear fit to time to peak likelihood for motion stimuli had a large negative intercept of -82ms (95% CIs: -124 to -40ms), which is substantially lower than the intercept for static trials. At the beginning of the time-window of interest (t_{train} = 75ms), the motion regression line crossed the y-axis at -1ms, while the static regression line crossed at 83ms. The mean distance between the two lines was 73ms, implying that the position of a moving object was represented with a latency that was approximately 70ms shorter than a static object in the same position. For early neural position representations (training times around 70-80ms), the latency of peak position representation was approximately 0ms. In turn, this means that these neural position representations were activated at the time that the moving object was physically centred on the corresponding position. Based on the training time, these representations likely originated in early visual cortex (V1-3), meaning that the early visual system was able to almost completely compensate for neural delays accumulated during processing up to that point and represent moving objects close to their real-time position.

3 Discussion

In this study we investigated how the visual system compensates for neural transmission delays when tracking the positions of moving objects. We investigated the latency of neural position representations of moving stimuli compared to unpredictably presented static stimuli, as well as the real-time position of the stimulus. By computing the timepoint at which each position representation of the moving stimulus was most similar to a static stimulus in the same location, we tracked the represented position of the stimulus over time across the visual hierarchy.

We demonstrate that classifiers trained to locate static stimuli based on the stimulus-evoked EEG signal could also localise moving stimuli. This is the first study to demonstrate cross-classification between stationary and smoothly moving stimuli with EEG, and gave us access to the fine temporal resolution needed to investigate the timing of neural responses to motion in humans. We subsequently showed that, during the first feedforward sweep of activity, the neural response encoding the position of a moving object was shifted approximately 70ms earlier than the response to a static stimulus. The early decoded representations of the position of a moving stimulus aligned with the real-time position of the object, rather than the position corresponding to afferent retinal input (subject to transmission and processing delays) which would instead signal outdated position information in visual cortex. Finally, we showed that delay compensation was primarily achieved before information reached visual cortex, as later processing of static and motion stimuli followed a similar time-course. Overall, this study shows the first direct neural evidence of motion extrapolation enabling accurate real-time representation of moving objects in humans.

These results are consistent with findings of receptive field (RF) shifts across the visual cortex in response to motion. Many earlier fMRI studies showing RF shifts against the direction of motion[30;31;32] have been dismissed because of the 'aperture-inward' bias, in which the trailing edge of a motion stimulus evokes larger responses than the leading edge[33]. This is not an issue for the present study, as we can determine the timing of neural responses at a fine temporal scale, rather than looking at aggregate responses over whole motion trajectories. Neural recordings from animals

and more recent fMRI studies in humans have reliably shown RF shifts throughout the visual cortex in response to motion[14], and that these displacements are against the direction of motion[13;7;34]. However, several differences remain between the previous fMRI results and the present study. Harvey and Dumoulin[14] found that RF shifts in response to motion scale with the size of the RF across the visual hierarchy. This implies that visual areas higher up the processing hierarchy that are activated later in time, for example MT, would shift their RFs more than lower visual areas, for example V1, which are activated earlier. In contrast, our results suggest that later visual areas show RF shifts of the same magnitude as earlier visual areas. However, it is not clear whether it is possible to map time elapsing after stimulus onset in EEG to processing in different visual areas as recorded by fMRI. The longer timescale of the fMRI signal means that it could be indexing later activity than we are recording with EEG, or include signals that emerge after integration of many feedforward and feedback sweeps of activity. While further research is needed to understand how extrapolation operates at different spatial scales, this converging evidence of RF shifts against the direction of motion suggests that the positions of moving objects are predictively encoded during processing, such that they are represented closer to their real-time position. Furthermore, we provide novel evidence that RF shifts likely correspond to changes in position representations during the initial feedforward sweep of the visual response.

Our findings point to several mechanisms that have been proposed to compensate for neural delays. We found that the early visual response to moving stimuli is shifted in time, such that the neural delays accumulated up to that point are compensated. However, during subsequent cortical processing, there is no further compensation for delays. As discussed in the Introduction, retinal ganglion cells respond to the leading edge of moving stimuli[4]. This effectively shifts the encoding of the position of a moving stimulus forward relative to a static stimulus at the earliest stage of processing. Additionally, evidence of a latency advantage for moving stimuli has been identified in the cat lateral geniculate nucleus of the thalamus[35], where visual information is transmitted en-route to the visual cortex. However, none of the previous evidence suggests that these pre-cortical mechanisms are sufficient to account for compensation for neural delays, to the extent we observe here. Therefore, it is likely that some cortical mechanisms do play a role. For example, there is evidence that a model of object motion is encoded in MT+, and influences neural response profiles in earlier visual areas through feedback connections[12;36]. These feedback connections could transmit information to neurons into whose receptive fields the moving stimulus will soon enter, driving an anticipatory response. Similarly, within-layer horizontal connections might activate neurons further ahead on the motion path[37]. Benvenuti et al.[37] show that this input from feedback and horizontal connections can drive spiking responses in cat V1. Crucially, our findings suggest that these mechanisms act *only* early in the course of stimulus processing, and therefore are present only early in the visual cortical hierarchy.

Even though we find temporal alignment between the early representations of the stimulus and its physical position, this alignment is lost during further processing. In a recent theoretical paper, Hogendoorn and Burkitt[15] argue that cortical motion extrapolation is necessary to minimise the discrepancy (prediction error) between an internal model of object position and the external world in the case of time-varying stimulation. There are two possible implementations of this cortical extrapolation: either delays are compensated through extrapolation in both feedforward and feedback activity, or, alternatively, extrapolation only occurs in feedback activity. Although the authors argue that the model including feedforward and feedback extrapolation is more parsimonious, this study suggests that feedforward cortical delays are not compensated. We therefore support the proposition that, if prediction errors are to be minimised, extrapolation might be implemented only in feedback connections. However, the present analysis approach may not be suitable to uncover this process, as cortical extrapolation could be a motion-specific computation enacted by different neural populations from those that encode static stimuli. Nevertheless, a complete model of compensation for neural delays in motion perception should account for extensive extrapolation early in visual processing, as observed here.

A limitation of the present study is that the localisation accuracy of moving stimuli was considerably lower than that of static stimuli. This is because classifiers were trained and tested on different stimulus types; neural populations that encode the position of static stimuli do not completely overlap with neural populations that encode the position of moving objects[38;6]. Additionally, previous fMRI studies show that, following a strong onset response, the neural

response to moving stimuli decreases over time[39;40;41;42]. This potentially leads to a decreasing signal-to-noise ratio over the course of a single motion trial. Furthermore, because motion epochs were quite long (up to 5s), the later parts of each motion trial could have been susceptible to slow drift of the EEG signal. Nevertheless, significant cross-classification between static and motion trials was still achieved, ruling this out as a major problem.

We additionally found that the earliest signals containing information about static stimuli were not informative about the location of the moving stimulus. Because of the spatial uncertainty associated with EEG, we do not know exactly where signals originate in the brain; source localisation in EEG is an ill-posed problem without co-registration with fMRI[43]. However, the timing of the earliest flash-evoked activity ($\sim$60ms after stimulus onset) suggests a source within V1[44;45;46;47;48]. In contrast, the position of the moving stimulus was decodable only on the basis of representations formed after approximately 100ms. The lack of cross-generalisation when training on early static stimulus-evoked activity suggests that this analysis approach does not capture the earliest motion-evoked V1 activity. One possibility is that, due to variability in when stimulus processing begins across trials[49], the signal-to-noise ratio in the static trials at these earlier timepoints may be too low to cross-generalise to moving stimuli. Alternatively, early processing of motion could be different to static stimuli: there is some evidence that direct connections from either LGN or the pulvinar to MT+ (bypassing V1) are used when processing motion[50]. This issue is hard to overcome, as training classifiers on moving stimuli would render capturing latency differences impossible; any latency shift in the test data would also be present in the training data. However, one promising approach was taken by Benvenuti et al.[37], who used recordings of monkey V1 to compare responses to trajectories of different lengths. They found that response latency decreased with increasing trajectory length: sub-threshold activation built up in front of the moving stimuli, preparing neural populations to fire upon the arrival of the stimulus in their RF. A similar approach could be taken in human EEG research to avoid the comparison between moving and non-moving stimuli. Additionally, this line of research would benefit from use of fMRI co-registered with EEG, which provides the temporal and spatial resolution necessary to pinpoint signals to a particular time, stimulus position and neural source.

Of relevance to these results is the flash-lag effect (FLE), a visual illusion in which a moving bar is perceived ahead of a flashed bar despite them being physically aligned[1]. This illusion demonstrates that moving objects are indeed perceived in an extrapolated position. Theories of the FLE can mainly be sorted into two camps: spatial explanations and temporal explanations[36]. Spatial models, for example motion extrapolation[1;51], suggest that the encoded positions of moving objects are shifted forwards to compensate for neural delays. In contrast, temporal models, for example differential latencies[52;53], suggest that motion is processed faster than flashes or that there is a temporal integration window over which position signals are averaged[54;55]. A range of psychophysical evidence has been presented to support each of these theories (and others), suggesting they all play a role in the FLE and, therefore, motion processing. However, our results are congruent only with spatial explanations; temporal models cannot explain how latency shifts could be greater than the latency of the unshifted neural response. We show that parts of the visual system encode moving objects at a position that afferent sensory information could not yet indicate. A similar result was found using EEG analysis of apparent motion[56], where a sensory template of an expected stimulus within the apparent motion sequence was pre-activated, before any sensory evidence was present. An outstanding question remains about whether neural representations of moving objects flexibly incorporate information about stimulus speed, as seen in animal V1 recordings[5;6] and the FLE[57].

3.1 Conclusion

This study used multivariate analysis of EEG data to investigate the latency of position representations of moving and static stimuli. We show that, during the first feedforward sweep of activity, the latency of the neural response to moving stimuli is substantially reduced compared to the response to unpredictable static stimuli. The effect of this latency advantage is that early visual areas represent moving objects in their real-time position, suggesting that (potentially a combination of) retinal, subcortical and cortical extrapolation mechanisms can overcome neural delays very early on in visual processing. Additional delays accumulated during subsequent cortical processing appear not to be compensated. These results demonstrate that the visual system predictively encodes the position of moving stimuli,

and provide an evidence base to constrain models of how and when motion extrapolation is achieved in the human visual system.

4 Acknowledgements

The authors gratefully acknowledge support from the Australian Research Council to HH (DP180102268 and FT200100246). This research was further supported by The University of Melbourne's Research Computing Services and the Petascale Campus Initiative. Thanks to Andrea Titton for discussion on calculating and reporting likelihoods, and to Jane Yook and Vinay Mepani for help with data collection.

5 Author Contributions

Conceptualisation: PJ & HH; Experiment Design: PJ & HH; Data Acquisition: PJ; Equipment: SB & HH; Programming: PJ; Analysis: PJ, TB & HH; Interpretation: PJ, TB, DF & HH; Writing - Original Draft: PJ; Writing - Review and Editing: PJ, TB, SvG, DF, SB & HH; Supervision: SvG, SB, & HH; Funding Acquisition: HH.

6 Conflict of Interest

The authors declare no competing financial interests.

7 Methods and Materials

7.1 Participants

Twelve participants (2 male; mean age = 27.0yrs, s.d. = 4.93yrs) completed all six testing sessions and were included in analyses. These were drawn from a larger initial pool of participants, including an additional fifteen participants that completed only the first session, which was used for screening. Of these additional participants, two withdrew from the study, three were excluded as the eyetracker could not consistently track their eye position, and the remaining ten were excluded after analysis of their first session data, due to poor fixation (more than 15% of trials with fixation lost) or poor EEG classification performance (less than 51.5% average classification accuracy when discriminating the location of static trials). Exclusion criteria included requiring glasses to view the computer screen and a personal or family history of epilepsy. Participants were recruited online through SONA and gave written informed consent before participation. Participants were reimbursed AU$15/hour for their time, as well as an additional AU$20 if they completed all six sessions. Ethical approval was granted by the University of Melbourne Ethics Committee (Ethics ID: 1954628.2).

7.2 Experimental Design

Stimuli were presented using MATLAB Version R2018a and the Psychophysics Toolbox extension version 3 [58;59;60]. Stimuli were presented on an ASUS ROG PG258 monitor (ASUS, Taipei, Taiwan) with a resolution of 1920 X 1080 running at a refresh rate of 200 Hz. Participants were seated, resting their heads on a chinrest, at a viewing distance of 50cm from the screen in a quiet, dimly-lit room.

Figure 1 shows the stimulus configuration and trial structure of the experiment. Stimuli were presented on a grey background, with a central fixation target [61]. Stimuli were black, filled circles with a radius of 1.29 degrees visual angle (dva) presented in a hexagonal configuration with 37 possible stimulus positions. A trial consisted of the stimulus flashing in a single location on the grid for 250ms (static trials), or moving in a straight line at a velocity of 10.36dva/s through the grid (motion trials), such that the amount of time spent travelling the length of the stimulus diameter was the same as the duration of the static stimulus. Motion vectors started and finished 7dva away from the grid to reduce

the effects of stimulus onset on the EEG signal. The stimulus passed through between four and seven flash positions, depending on the eccentricity of the vector, taking 400ms to travel between grid positions. Static and motion trials were randomly shuffled within each experimental session, with an inter-stimulus interval randomly selected from a uniform distribution between 350 and 450ms. In each testing session, each static stimulus location was repeated 42 times, while each of the 42 motion vectors (6 directions X 7 parallel starting positions) was repeated 18 times. Trials were split into seven blocks, with a duration of approximately nine minutes each. After each block, participants could rest and sit back from the chinrest. Six times within each block (every 50 trials), participants could take a mini-break, in which the experiment was paused but they were required to remain in the chinrest. This procedure was repeated over six sessions, totalling 252 static trials in each location and 108 repetitions of each motion vector.

Participants performed a simple target detection task in order to ensure they attended to the stimuli. While maintaining fixation on the fixation point at the centre of the screen, they responded as quickly as possible with the space-bar when the stimulus flashed red for 40ms. This happened at random 45 times per block, and trials containing a target were discarded from analysis to ensure that the target and response did not interfere with the relevant EEG analysis. Each of the target trials was then repeated at the end of the block without a target to maintain equal trial numbers for each static stimulus position/motion vector. Participants completed one practice block of twenty trials at the start the first session to become acquainted with the task. The practice block could be repeated upon request.

EEG and eyetracking data were collected from participants while they viewed the stimuli. Eyetracking data were collected using an EyeLink 1000 eye tracker (SR Research). The eyetracker was calibrated at the start of each block, and drift correction was applied after each mini-break. The conversion of the EyeLink 1000 *.edf* files to *.mat* files and offline fixation checks were performed with the EyeCatch toolbox [62].

Continuous EEG data were recorded at 2048Hz using a 64-channel BioSemi Active-Two system (BioSemi, Amsterdam, The Netherlands), connected to a standard 64-electrode EEG cap. Two external electrodes were placed on the mastoids, to be used as a reference. Electrooculography (EOG) was recorded using six electrodes: on the canthi (horizontal) and above and below the eyes (vertical).

7.3 EEG Pre-processing

EEG pre-processing was conducted using EEGLAB version 2021.1 [63], running in MATLAB R2017b. First, EEG data were re-referenced to the mastoid channels. Data were down-sampled to 128Hz to reduce computation time and memory load required for further pre-processing and analysis. No filtering was applied to data so as not to distort event timing [64]. Bad channels were noted during data collection and were interpolated using spherical interpolation. On average, 0.49 electrodes were interpolated per recording session. Additionally, one complete session was dropped from further analysis for one participant, due to a poor connection to the mastoid channels. Data were epoched from 100ms before flash/motion onset to 100ms after flash/motion offset. The 100ms period before onset was used to baseline correct each epoch, by subtracting the mean amplitude in this period from the whole epoch.

Eye movement data were used to check fixation: static trials in which gaze deviated more than 2.1dva from fixation (i.e. was closer to another stimulus position than the central fixation point) at any point while the stimulus was on screen were discarded from analysis, as these eye movements would disrupt retinotopy. On average, 11.2% of trials were rejected on this basis. Participants' eye positions during flashes were further analysed to ensure that there were no systematic eye movements which could be exploited by classifiers during the EEG analysis (see Supplementary Figure 1). No motion trials were rejected on the basis of eye movements. This is because motion trials were only used for testing classifiers; if no systematic eye movements are present in the training set, then the classifier cannot learn to distinguish trials on the basis of eye movements, so any eye movements in the test data are irrelevant to the analysis.

Epochs were then automatically rejected through an amplitude threshold. For static trials, epochs were rejected if the standard deviation of the amplitude of any channel exceeded four standard deviations from the mean standard deviation of that channel across all epochs. This resulted in 8.3% of epochs being rejected across all observers. Motion trials were rejected with a threshold of five standard deviations from the mean standard deviation. This less

stringent threshold reflects the longer duration of motion trials; more variability in amplitude can be expected. 7.5% of motion trials were rejected across all observers. Finally, static and motion epochs were demeaned. The average amplitude of each electrode across all static stimulus locations was subtracted from each trial amplitude, while for motion trials, the average amplitude from motion vectors of the same length was subtracted. This ensured that the classifiers could leverage any changes in the signal corresponding to stimulus location, without the potential confound of overall amplitude differences in static compared to motion trials[22].

7.4 EEG Analysis

Analyses were programmed using MATLAB Version R2017b and run on the University of Melbourne Spartan High Performance Computing system. Time-resolved multivariate pattern analysis was used to classify EEG data according to the location of the static stimuli. Linear discriminant analysis (LDA) classifiers with a shrinkage regularisation parameter of 0.05[65] were trained to discriminate the location of static stimuli at timepoints from 0-350ms (i.e. from stimulus onset to 100ms after stimulus offset). Code for classification analysis was adapted from Mostert et al.[65] and Hogendoorn and Burkitt[18].

In this analysis, time elapsing post stimulus onset can be seen as a proxy for processing stage. As time passes, stimulus-evoked activity will progress through the visual system[22;66]. Our aim was to establish, for each training timepoint, the timepoint in the test data at which the stimulus was most likely to be at a certain position. This tells us the latency of a particular pattern of activity, or representation of the stimulus, in the training data compared to the test data.

We first demonstrated that stimulus position could be discriminated even when static stimuli were close together, by averaging classification results according to distance between stimulus locations. Next, we calculated the latency of representations when training and testing on static stimuli. This was used as a baseline to which the motion was compared, as the static stimulus locations were unpredictable. Any shifts in latency seen in the motion trials must be due to the predictable preceding trajectory. The key analysis was, therefore, training classifiers to discriminate the location of static stimuli and testing on motion vectors. In this case, in the training data, the stimulus was centred at a certain position, so the timepoint at which the test data is most similar should be the timepoint at which the stimulus was represented in the brain at this position in the trajectory. This analysis was repeated across all training times, which allowed us to 'track' the neural representation of the stimulus as it was processed and compare compensation for neural delays at different stages of processing. Full compensation for neural delays would be seen if the brain represents moving stimuli aligned to their real-time position, while if there was no compensation, there would be no difference in the latency of the peak response between static and moving stimuli.

To avoid bias that often emerges from multi-class classification[67], classifiers were trained using pairwise combinations of stimulus positions, such that a classifier was trained to discriminate each location from every other location. As it is redundant to train classifiers to discriminate e.g., position 1 vs 2 and also 2 vs 1, this resulted in in 666 trained classifiers at 90 timepoints over a 350ms period. The number of trials in each class was balanced by sampling trials without replacement from the majority class to equal the number of trials in the minority class. These classifiers were then tested on either unseen static trials (five-fold cross-validation between train and test sets) or motion trials. At each timepoint, pairwise classification results were combined to estimate the likelihood of the stimulus being in a given position[68;69]. We can estimate $\mathbb{P}(\text{position } i|\text{stimulus is in } s)$ as

$$p_s(i) = \left(\sum_{j \neq i} \frac{1}{r_{ij}} - (k-2) \right)^{-1}$$

where r_{ij} is the classification performance for position i vs position j, and k is the total number of classes (the 37 stimulus positions, in this case). Such that the probability across all positions was equal to 1, the estimated likelihoods were then normalised between 0 and 1. If decoding performance was at chance-level, we would expect uniform

likelihood across all stimulus positions, at:

$$\frac{1}{k} \approx 0.027027\ldots$$

A likelihood greater than this indicates a location-specific neural response to the stimulus. An example probability mass functions of the likelihood across all stimulus positions can be found in Figure 3 step 2. For the main analysis, we investigated the evolution over time of the likelihood of the stimulus being at the presented position, $p_s(s)$, referred to as the stimulus-position likelihood. Where relevant, chance-level (1/37) was subtracted from the likelihood for easier interpretation in graphs.

7.5 Statistical Analysis

Statistical significance of classification results was ascertained through permutation testing. After running the classification analyses as described above, class labels were randomly shuffled when calculating the stimulus-position likelihood, ensuring that the permuted classification results were uninformative about stimulus location. This procedure was repeated 1000 times per participant, providing a null distribution against which our results could be compared with Yuen's t-test, one-tailed, $\alpha = 0.05$[70]. Cluster-based correction for multiple comparisons was applied with 1000 permutations (cluster-forming alpha = 0.05,[71;72]). Code for the cluster-based correction came from the Decision Decoding Toolbox[73], which uses code originally from LIMO EEG[74] to implement Yuen's t-test.

To test significance of linear regression models against a constant model, we used one-tailed *F*-tests. To test whether individual regression coefficients were significantly different from zero, we used two-tailed *t*-test.

7.6 Data Availability

Code and summarised data files, which can be used to reproduce all figures, will be made available at `https://osf.io/jbw9m/`. Raw or pre-processed data files are available upon request, to reduce the environmental impact of hosting large datafiles online unnecessarily[75].

References

[1] Nijhawan, R. Motion extrapolation in catching. *Nature* **370**, 256–257 (1994). URL `http://www.nature.com/articles/370256b0`. Publisher: Nature Publishing Group.

[2] Friston, K. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience* **11**, 127–138 (2010).

[3] Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* **36**, 181–204 (2013). Publisher: Cambridge University Press.

[4] Berry, M. J., Brivanlou, I. H., Jordan, T. A. & Meister, M. Anticipation of moving stimuli by the retina. *Nature* **398**, 334–338 (1999).

[5] Jancke, D., Erlhagen, W., Schöner, G. & Dinse, H. R. Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex. *J Physiol* **556**, 971–982 (2004).

[6] Subramaniyan, M. *et al.* Faster processing of moving compared with flashed bars in awake macaque V1 provides a neural correlate of the flash lag illusion. *Journal of Neurophysiology* **120**, 2430–2452 (2018).

[7] Sundberg, K. A., Fallah, M. & Reynolds, J. H. A motion-dependent distortion of retinotopy in area V4. *Neuron* **49**, 447–457 (2006). URL `http://www.cell.com/article/S0896627305011360/fulltext`. Publisher: Elsevier.

[8] Chen, E. Y. *et al.* Alert Response to Motion Onset in the Retina. *The Journal of Neuroscience* **33**, 120–132 (2013).

[9] Johnston, J. & Lagnado, L. General features of the retinal connectome determine the computation of motion anticipation. *eLife* **2015** (2015). Publisher: eLife Sciences Publications Ltd.

[10] Liu, B., Hong, A., Rieke, F. & Manookin, M. B. Predictive encoding of motion begins in the primate retina. *Nature Neuroscience* **24** (2021).

[11] Souihel, S. & Cessac, B. On the potential role of lateral connectivity in retinal anticipation. *The Journal of Mathematical Neuroscience* **11** (2020). URL http://arxiv.org/abs/2009.02081.

[12] Maus, G. W., Fischer, J. & Whitney, D. Motion-dependent representation of space in area MT+. *Neuron* **78**, 554–562 (2013).

[13] Schneider, M. *et al.* Motion Displaces Population Receptive Fields in the Direction Opposite to Motion. *bioRxiv* 759183 (2019). URL http://dx.doi.org/10.1101/759183. Publisher: Cold Spring Harbor Laboratory.

[14] Harvey, B. M. & Dumoulin, S. O. Visual motion transforms visual space representations similarly throughout the human visual hierarchy. *NeuroImage* **127**, 173–185 (2016). URL http://dx.doi.org/10.1016/j.neuroimage.2015.11.070.

[15] Hogendoorn, H. & Burkitt, A. N. Predictive coding with neural transmission delays: a real-time temporal alignment hypothesis. *eNeuro* **6** (2019). URL http://dx.doi.org/10.1101/453183.

[16] van Heusden, E., Harris, A. M., Garrido, M. I. & Hogendoorn, H. Predictive coding of visual motion in both monocular and binocular human visual processing. *Journal of Vision* **19**, 3 (2019). URL http://jov.arvojournals.org/article.aspx?doi=10.1167/19.1.3. Publisher: The Association for Research in Vision and Ophthalmology.

[17] King, J.-R. & Dehaene, S. Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in cognitive sciences* **18**, 203–210 (2014). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5635958/.

[18] Hogendoorn, H. & Burkitt, A. N. Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. *NeuroImage* **171**, 55–61 (2018). URL https://doi.org/10.1016/j.neuroimage.2017.12.063.

[19] Haufe, S. *et al.* On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* **87**, 96–110 (2014).

[20] Johnson, P., Grootswagers, T., Moran, C. & Hogendoorn, H. Temporal dynamics of visual population receptive fields. In *43rd European Conference on Visual Perception (ECVP) 2021 Online*, vol. 50, 1–244 (SAGE Publications Ltd STM, 2021). URL https://doi.org/10.1177/03010066211059887.

[21] Harvey, B. M. & Dumoulin, S. O. The relationship between cortical magnification factor and population receptive field size in human visual cortex: Constancies in cortical architecture. *Journal of Neuroscience* **31**, 13604–13612 (2011). URL http://white.stanford.edu/software/.

[22] Dijkstra, N., Ambrogioni, L., Vidaurre, D. & van Gerven, M. Neural dynamics of perceptual inference and its reversal during imagery. *eLife* **9** (2020).

[23] Sack, A. T., Kohler, A., Linden, D. E., Goebel, R. & Muckli, L. The temporal characteristics of motion processing in hMT/V5+: Combining fMRI and neuronavigated TMS. *NeuroImage* (2006).

[24] Ahlfors, S. P. *et al.* Spatiotemporal Activity of a Cortical Network for Processing Visual Motion Revealed by MEG and fMRI. *Journal of Neurophysiology* **82**, 2545–2555 (1999). URL https://journals.physiology.org/doi/full/10.1152/jn.1999.82.5.2545.

[25] Hotson, J., Braun, D., Herzberg, W. & Boman, D. Transcranial magnetic stimulation of extrastriate cortex degrades human motion direction discrimination. *Vision Research* **34**, 2115–2123 (1994). URL https://www.sciencedirect.com/science/article/pii/0042698994903212.

[26] Yoshor, D., Bosking, W. H., Ghose, G. M. & Maunsell, J. H. R. Receptive Fields in Human Visual Cortex Mapped with Surface Electrodes. *Cerebral Cortex* **17**, 2293–2302 (2007). URL https://academic.oup.com/cercor/article/17/10/2293/310030. Publisher: Oxford Academic.

[27] Mohsenzadeh, Y., Qin, S., Cichy, R. M. & Pantazis, D. Ultra-rapid serial visual presentation reveals dynamics of feedforward and feedback processes in the ventral visual pathway. *eLife* **7** (2018). Publisher: eLife Sciences Publications Ltd.

[28] Lamme, V. A. F., Roelfsema, P. R. & Roelfsma, P. R. The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences* **23**, 571–579 (2000). Publisher: Elsevier Current Trends.

[29] D'Errico, J. SLM - Shape Language Modeling, MATLAB Central File Exchange. (2022). URL `https://www.mathworks.com/matlabcentral/fileexchange/24443-slm-shape-language-modeling`.

[30] Liu, J. V., Ashida, H., Smith, A. T. & Wandell, B. A. Assessment of Stimulus-Induced Changes in Human V1 Visual Field Maps. *Journal of Neurophysiology* **96**, 3398–3408 (2006). URL `https://www.physiology.org/doi/10.1152/jn.00556.2006`. Publisher: American Physiological Society.

[31] Whitney, D. *et al.* Flexible Retinotopy: Motion-Dependent Position Coding in the Visual Cortex. *Science* **302**, 878–881 (2003). URL `https://science.sciencemag.org/content/302/5646/878`.

[32] Raemaekers, M., Lankheet, M. J., Moorman, S., Kourtzi, Z. & van Wezel, R. J. Directional anisotropy of motion responses in retinotopic cortex. *Human Brain Mapping* **30**, 3970–3980 (2009). URL `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6870668/`.

[33] Wang, H. X., Merriam, E. P., Freeman, J. & Heeger, D. J. Motion direction biases and decoding in human visual cortex. *Journal of Neuroscience* **34**, 12601–12615 (2014). Publisher: Society for Neuroscience.

[34] Fu, Y. X., Shen, Y., Gao, H. & Dan, Y. Asymmetry in Visual Cortical Circuits Underlying Motion-Induced Perceptual Mislocalization. *Journal of Neuroscience* **24**, 2165–2171 (2004). URL `www.jneurosci.org`. Publisher: Society for Neuroscience.

[35] Orban, G. A., Hoffmann, K. P. & Duysens, J. Velocity selectivity in the cat visual system. I. Responses of LGN cells to moving bar stimuli: a comparison with cortical areas 17 and 18. *Journal of Neurophysiology* **54**, 1026–1049 (1985). URL `https://journals.physiology.org/doi/abs/10.1152/jn.1985.54.4.1026`. Publisher: American Physiological Society.

[36] Maus, G. W., Ward, J., Nijhawan, R. & Whitney, D. The Perceived Position of Moving Objects: Transcranial Magnetic Stimulation of Area MT+ Reduces the Flash-Lag Effect. *Cerebral Cortex* **23**, 241–247 (2013). URL `https://academic.oup.com/cercor/article-lookup/doi/10.1093/cercor/bhs021`.

[37] Benvenuti, G. *et al.* Anticipatory responses along motion trajectories in awake monkey area V1. *BioRXiv* (2020). URL `https://www.biorxiv.org/content/10.1101/2020.03.26.010017v2`.

[38] Noda, H., Freeman, R. B., Gies, B. & Creutzfeldt, O. D. Neuronal responses in the visual cortex of awake cats to stationary and moving targets. *Experimental Brain Research* **12**, 389–405 (1971). URL `https://doi.org/10.1007/BF00234494`.

[39] Schellekens, W. *et al.* Predictive coding for motion stimuli in human early visual cortex. *Brain Structure and Function* **221**, 879–890 (2016). URL `http://link.springer.com/10.1007/s00429-014-0942-2`. Publisher: Springer Berlin Heidelberg.

[40] Schellekens, W., Ramsey, N. F., van Wezel, R. J. & Raemaekers, M. Changes in fMRI BOLD dynamics reflect anticipation to moving objects. *NeuroImage* **161**, 188–195 (2017). URL `https://www-sciencedirect-com.ezp.lib.unimelb.edu.au/science/article/pii/S1053811916304797`.

[41] Harrison, L. M., Stephan, K. E., Rees, G. & Friston, K. J. Extra-classical receptive field effects measured in striate cortex with fMRI. *NeuroImage* **34**, 1199–1208 (2007). Publisher: Academic Press.

[42] McKeefry, D. J., Watson, J. D., Frackowiak, R. S., Fong, K. & Zeki, S. The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion. *NeuroImage* **5**, 1–12 (1997). Publisher: Academic Press Inc.

[43] Jatoi, M. A., Kamel, N., Malik, A. S., Faye, I. & Begum, T. A survey of methods used for source localization using EEG signals. *Biomedical Signal Processing and Control* **11**, 42–52 (2014). URL https://www.sciencedirect.com/science/article/pii/S174680941400010X.

[44] Alilović, J., Timmermans, B., Reteig, L. C., van Gaal, S. & Slagter, H. A. No Evidence that Predictions and Attention Modulate the First Feedforward Sweep of Cortical Information Processing. *Cerebral Cortex* **29**, 2261–2278 (2019). URL https://academic.oup.com/cercor/advance-article/doi/10.1093/cercor/bhz038/5382202. Publisher: Oxford University Press.

[45] Fahrenfort, J. J., Scholte, H. S. & Lamme, V. A. F. Masking disrupts recurrent processing in human visual cortex. *Journal of Cognitive Neuroscience* **19**, 1488–1497 (2009).

[46] Di Russo, F., Martinez, A., Sereno, M. I., Pitzalis, S. & Hillyard, S. A. Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping* **15**, 95–111 (2002). URL https://onlinelibrary.wiley.com/doi/10.1002/hbm.10010.

[47] Wibral, M., Bledowski, C., Kohler, A., Singer, W. & Muckli, L. The timing of feedback to early visual cortex in the perception of long-range apparent motion. *Cerebral Cortex* **19**, 1567–1582 (2009). URL https://academic.oup.com/cercor/article/19/7/1567/316530. Publisher: Oxford Academic.

[48] Vanni, S. *et al.* Timing of interactions across the visual field in the human cortex. *NeuroImage* **21**, 818–828 (2004). Publisher: Academic Press.

[49] Vidaurre, D., Myers, N. E., Stokes, M., Nobre, A. C. & Woolrich, M. W. Temporally Unconstrained Decoding Reveals Consistent but Time-Varying Stages of Stimulus Processing. *Cerebral Cortex* **29**, 863–874 (2019). URL https://doi.org/10.1093/cercor/bhy290.

[50] Ffytche, D. H., Guy, C. N. & Zeki, S. The parallel visual motion inputs into areas V1 and V5 of human cerebral cortex. *Brain* **118**, 1375–1394 (1995). URL https://academic.oup.com/brain/article/118/6/1375/319377. Publisher: Oxford Academic.

[51] Hogendoorn, H. Motion Extrapolation in Visual Processing: Lessons from 25 Years of Flash-Lag Debate. *Journal of Neuroscience* **40**, 5698–5705 (2020). URL https://www.jneurosci.org/content/40/30/5698.

[52] Whitney, D. & Murakami, I. Latency difference, not spatial extrapolation. Tech. Rep., Plenum (1998). URL http://neurosci.nature.com. Publication Title: Nevo, E. in Evolutionary Biology Volume: 387 Issue: 9.

[53] Whitney, D. & Cavanagh, P. Motion distorts visual space: Shifting the perceived position of remote stationary objects. *Nature Neuroscience* **3**, 954–959 (2000). Publisher: Nature Publishing Group.

[54] Krekelberg, B. & Lappe, M. The position of moving objects. *Science* **289**, 1107a–1107 (2000). Publisher: American Association for the Advancement of Science (AAAS).

[55] Whitney, D., Murakami, I. & Cavanagh, P. Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research* **40**, 137–149 (2000). URL https://www.sciencedirect.com/science/article/pii/S0042698999001662. Publisher: Pergamon.

[56] Blom, T., Feuerriegel, D., Johnson, P., Bode, S. & Hogendoorn, H. Predictions drive neural representations of visual events ahead of incoming sensory information. *Proceedings of the National Academy of Sciences of the United States of America* **117** (2020).

[57] Wojtach, W. T., Sung, K., Truong, S. & Purves, D. An empirical explanation of the flash-lag effect. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 16338–16343 (2008).

[58] Brainard, D. H. The psychophysics toolbox. *Spatial Vision* **10**, 433–436 (1997).

[59] Pelli, D. G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision* **10**, 437–442 (1997).

[60] Kleiner, M., Brainard, D. & Pelli, D. "What's new in Psychtoolbox-3?". In *Perception 36 ECVP Abstract Supplement* (2007).

[61] Thaler, L., Schütz, A. C., Goodale, M. A. & Gegenfurtner, K. R. What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research* **76**, 31–42 (2013). Publisher: Pergamon.

[62] Bigdely-Shamlo, N., Kreutz-Delgado, K., Kothe, C. & Makeig, S. EyeCatch: Data-mining over Half a Million EEG Independent Components to Construct a Fully-Automated Eye-Component Detector. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* **2013**, 5845–5848 (2013). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4136453/.

[63] Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* **134**, 9–21 (2004).

[64] van Driel, J., Olivers, C. N. L. & Fahrenfort, J. J. High-pass filtering artifacts in multivariate classification of neural time series data. *bioRxiv* 530220 (2019). URL http://dx.doi.org/10.1101/530220. Publisher: Cold Spring Harbor Laboratory.

[65] Mostert, P., Kok, P. & De Lange, F. P. Dissociating sensory from decision processes in human perceptual decision making. *Scientific Reports* **5**, 18253 (2015). URL www.nature.com/scientificreports. Publisher: Nature Publishing Group.

[66] King, J.-R. & Wyart, V. The Human Brain Encodes a Chronicle of Visual Events at each Instant of Time thanks to the Multiplexing of Traveling Waves. *The Journal of Neuroscience* (2021).

[67] Yan, C. *et al.* Self-weighted Robust LDA for Multiclass Classification with Edge Classes. *ACM Transactions on Intelligent Systems and Technology* **12**, 4:1–4:19 (2020). URL https://doi.org/10.1145/3418284.

[68] Manyakov, N. V. & Van Hulle, M. M. Decoding grating orientation from microelectrode array recordings in monkey cortical area v4. *International Journal of Neural Systems* **20**, 95–108 (2010). URL https://www.worldscientific.com/doi/abs/10.1142/S0129065710002280.

[69] Price, D., Knerr, S., Personnaz, L. & Dreyfus, G. Pairwise neural network classifiers with probabilistic outputs. In *Neural Information Processing Systems*, 1109–1116 (The MIT Press, Cambridge, MA, 1995).

[70] Yuen, K. K. The two-sample trimmed t for unequal population variances. *Biometrika* **61**, 165–170 (1974). URL https://doi.org/10.1093/biomet/61.1.165.

[71] Bullmore, E. *et al.* Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging* **18**, 32–42 (1999).

[72] Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods* **164**, 177–190 (2007). URL https://www.sciencedirect.com/science/article/pii/S0165027007001707.

[73] Bode, S., Feuerriegel, D., Bennett, D. & Alday, P. M. The Decision Decoding ToolBOX (DDTBOX) – A Multivariate Pattern Analysis Toolbox for Event-Related Potentials. *Neuroinformatics* **17**, 27–42 (2019). URL https://doi.org/10.1007/s12021-018-9375-z.

[74] Pernet, C. R., Chauveau, N., Gaspar, C. & Rousselet, G. A. LIMO EEG: A Toolbox for Hierarchical LInear MOdeling of ElectroEncephaloGraphic Data. *Computational Intelligence and Neuroscience* **2011**, e831409 (2011). URL https://www.hindawi.com/journals/cin/2011/831409/. Publisher: Hindawi.

[75] Rae, C. L., Farley, M., Jeffery, K. J. & Urai, A. E. Climate crisis and ecological emergency: Why they concern (neuro)scientists, and what we can do. *Brain and Neuroscience Advances* **6**, 23982128221075430 (2022). URL https://doi.org/10.1177/23982128221075430. Publisher: SAGE Publications Ltd STM.