Electrophysiological indicators of gesture perception

Maria E. Cabrera · Keisha Novak · Dan Foti · Richard Voyles · Juan P. Wachs

Abstract
Electroencephalography (EEG) activity in the mu frequency band (8–13 Hz) is suppressed during both gesture performance and observation. However, it is not clear if or how particular characteristics within the kinematic execution of gestures map onto dynamic changes in mu activity. Mapping the time course of gesture kinematics onto that of mu activity could help understand which aspects of gestures capture attention and aid in the classification of communicative intent. In this work, we test whether the timing of inflection points within gesture kinematics predicts the occurrence of oscillatory mu activity during passive gesture observation. The timing for salient features of performed gestures in video stimuli was determined by isolating inflection points in the hands’ motion trajectories. Participants passively viewed the gesture videos while continuous EEG data was collected. We used wavelet analysis to extract mu oscillations at 11 Hz and at central electrodes and occipital electrodes. We used linear regression to test for associations between the timing of inflection points in motion trajectories and mu oscillations that generalized across gesture stimuli. Separately, we also tested whether inflection point occurrences evoked mu/alpha responses that generalized across participants. Across all gestures and inflection points, and pooled across participants, peaks in 11 Hz EEG waveforms were detected 465 and 535 ms after inflection points at occipital and central electrodes, respectively. A regression model showed that inflection points in the motion trajectories strongly predicted subsequent mu oscillations ($R^2 = 0.921, p<0.01$); effects were weaker and non-significant for low (17 Hz) and high (21 Hz) beta activity. When segmented by inflection point occurrence rather than stimulus onset and testing participants as a random effect, inflection points evoked mu and beta activity from 308 to 364 ms at central electrodes, and broad activity from 226 to 800 ms at occipital electrodes. The results suggest that inflection points in gesture trajectories elicit coordinated activity in the visual and motor cortices, with prominent activity in the mu/alpha frequency band and extending into the beta frequency band. The time course of activity indicates that visual processing drives subsequent activity in the motor cortex during gesture processing, with a lag of approximately 80 ms.

Keywords Gesture processing · Mirror neuron · Mu · EEG

Introduction
Performing gestures is a form of communication that is used to complement, emphasize, or accompany verbal messages, or alternatively serve as an integral modality of communication on its own – without the verbal component. The meaning of a gesture may differ across cultures and societies, yet the perception and execution of gestures are intrinsic parts of human behavior; as such, there are cognitive processes associated with gesture perception, production, and recognition. Relationships between subjective perception and production of an action are thought to be reflected by the human mirroring system (Decety and Grèzes 1999) activating regions of the brain involved in performing an action even when merely observed (Gazzola and Keysers 2009). These activations are
associated with facilitating interpretation and understanding of movements performed by others (Rizzolatti and Sinigaglia 2010). Specifically, empathy, communication, social skills, and overall coordination can all be traced back to the importance of motor representations (Urgen et al. 2013).

A highly relevant and salient finding to emerge from research on action perception is that both execution and observation of bodily movements activate the human Mirror Neuron System (hMNS; Di Pellegrino et al. 1992; Iacoboni et al. 1999; Raos et al. 2006; Rizzolatti and Craighero 2004). A prominent theory regarding the function of these neurons posits that the hMNS facilitates action understanding (Gallese and Sinigaglia 2012; Rizzolatti et al. 1996). Thus, these neurons exhibiting close correspondence between observed action and performed action are found in several cortical areas, including the premotor cortex and parietal areas (Iacoboni et al. 1999). Fronto-parietal hMNS activation is specific to type of action performed, including executed, imagined, and observed actions (Filimon et al. 2007), and that adaptation across action and perception demonstrates distinct areas of the brain that respond selectively to motoric and perceptual representations of actions (Chong et al. 2008). Additionally, hMNS have been ascribed a wide variety of psychological and perceptual functions including imitation (Iacoboni et al. 1999), language (Rizzolatti and Arbib 1998), empathy (Avenanti et al. 2005), emotion regulation (Enticott et al. 2008), and speech among others (Arbib 2005; Glenberg et al. 2008; Hobson and Bishop 2016; Kühn and Brass 2008; Théoret and Pascual-Leone 2002).

Though widely believed to be an adaptation for psychological function and comprehension of actions, the function of the hMNS has also been argued to reflect general associated learning processes in development. This “associative account” allows, but does not assume, that the hMNS makes positive contributions to social understanding (Cook et al. 2014). Specifically, instances have been observed where: (1) hMNS activation does not consistently encode action intention, (2) hMNS activation can be changed via sensorimotor training, and (3) activations of the hMNS may be contingent or context-sensitive (Cook et al. 2014). Thus, while the precise function of the hMNS is still under debate, it is clear that the human motor cortex subserves bodily action as well as mental representations thereof and is involved in gesture processing.

In recent years, there has been an increase in studies utilizing electroencephalography (EEG) to examine mu rhythm as a potential index of hMNS functioning. EEG offers precise temporal resolution of electrical activity at the scalp. Of particular relevance to gesture processing are mu rhythms, which occur within the alpha frequency band (8–13 Hz) and at central scalp electrodes, directly above the motor and premotor cortices. Decreases in mu power (i.e., mu desynchronization) are thought to reflect motor cortex activation not only during action execution but also observation (Braadbaart et al. 2013; Muthukumaraswamy et al. 2004). It has also been argued to play a role in contexts involving social interaction (Perry et al. 2011), and passive action observation (Hogeveen et al. 2014). In the same alpha frequency band, EEG activity at occipital electrodes has been associated with activation of the visual cortex and visual processing (Neuper et al. 2005; Pfurtscheller et al. 1994). Thus, while mu activity at central electrodes may capture activation of the hMNS during gesture processing, activity in the same frequency band at occipital electrodes captures visual processing more generally. For example, Quandt et al. (2012) showed that gesture performance had an effect on alpha oscillations through recorded EEG. Their study included different gesture types, such as deictic (i.e., pointing gestures) and iconic (i.e., pantomime, related to speech). Their results suggest that different types of gestures engage the motor system in different ways based on the pattern of alpha activity at central, parietal, and occipital sites. Quandt et al. (2012) found decreased power in alpha activity when observing iconic but not deictic gestures, suggesting that iconic gestures may elicit greater motor system activation due to the dynamic and richly descriptive motion of iconic gestures, compared to relatively static components of deictic gestures. Furthermore, Wu and Coulson (2005) used EEG and event-related potentials (ERPs) to examine neural components of iconic gestures. Their findings suggest iconic gestures elicit neural processes analogous to those evoked by other meaningful representations such as pictures and words. Thus, gesture observation elicits semantic processing that can be meaningfully captured by electrophysiological measures.

Complementary research using functional magnetic resonance imaging (fMRI) has examined whether it is simply the movement trajectory (a visual feature) or the task goal (the intention of the observer) that is crucial for hMNS activation (Engel et al. 2008). Their findings suggest that it is not a specific visual feature that activates the hMNS, rather that the hMNS seems to respond when an observed movement is matched to a motor representation triggered by the intentional goal of the observer (Engel et al. 2008). This aspect is clearly distinct from purely processing visual features. Such a correlation to movement execution can be constructed from both biological (performed by humans) and non-biological (i.e., robotic arms or virtual avatars) stimuli and it is conceivable that this may often arise implicitly or subconsciously, when visual features trigger strong associations with human movements. Relatively, Dahan and Reiner (2016) examined gesture motions performed by humans vs non-humans (artificial objects), and meaningful vs meaningless gestures, to find that the hMNS is only activated to motions performed by the kinematics of the human body. However, the part of the stimulus which triggers activation...
of the hMNS remains unclear: the limb-like Gestalt of the moving device, movement trajectories that resemble human-performed or biological trajectories, or, as suggested by Gazzola et al. (2007), comprehending the goal of an observed action even when performed by a robotic arm. One reason could be that the trajectories of the objects were processed as biological movements, because they followed a smooth path.

Whereas most previous studies have examined average hMNS activation during gesture processing, one study also examined the time course of neural activity. Specifically, the authors compared central mu activity and occipital alpha activity observed across different gesture types and across three time windows: 0–1 s, 1–2 s, and 2–3 s (Streltsova et al. 2010). They reported decreased central mu activity, but not occipital alpha activity, when the gestures observed were meaningless and not communicative. Critically, they also found differences in the time course, whereby meaningless gestures elicited a sustained reduction in mu activity, whereas communicative gestures exhibited significant fluctuations across the three time windows. This suggests that fluctuations in mu activity while observing communicative gestures is not necessarily related to the total amount of movement, but instead may indicate semantic processing and comprehension of gesture meaning. A limitation of that study, however, is that the time bins were relatively broad and were the same across all gestures, leaving it unclear which aspects of the gestures may be elicited fluctuations in mu activity.

Building on this finding, the current study tested whether the timing of mu oscillations during gesture observation would follow a unique time course for each gesture and could be mapped onto salient characteristics of gesture kinematics. To extract the time course of mu oscillations in a more fine-grained manner, the current study measured mu power continuously rather than in discrete time bins; peaks in mu activity were then used to infer the occurrence of salient gesture characteristics. Our analyses were largely exploratory, examining whether mu oscillations would correspond specifically with inflection points in the kinematic trace of the gesture manifestation. If so, this would suggest that inflection points are used as representations of salient aspects of gestures. That is, it is possible that gestures with similar inflection points may be perceived as similar regarding their meaning, whereas those with different inflection points may be perceived as distinct. We do not suggest that inflection points are the only salient characteristics of gesture kinematics, but rather they represent a reasonable starting point for unpacking the dynamics of gesture processing.

The purpose of this work is to examine how particular characteristics within the kinematic execution of gestures map onto dynamic changes in mu activity. Mapping the time course of gesture kinematics onto the time course of mu activity could help understand which aspects of gestures capture attention and aid in the classification of communicative intent. Our aim is to analyze neural signatures from EEG data recorded while observing gestures, to examine the relationship between key motion components of gestures with oscillations in the mu frequency band. The present work is an extension of work by Cabrera et al. (2017), in which a preliminary analysis was conducted on a subset of the current sample to determine the existence of a relationship between salient kinematic points in observed gestures with oscillations in EEG frequencies associated to activation in both motor and visual cortices. The current work focuses on corroborating these results in a larger sample and with additional analyses, with two specific aims:

1. **Map the timing of inflection points in gesture kinematics onto oscillatory EEG activity** We extracted the timing of each inflection point across a series of naturalistic gestures. Separately, we extracted the timing of local maxima in EEG activity in the mu/alpha frequency band (i.e., 11 Hz). Of interest was whether the occurrence in an inflection point would predict a subsequent peak in mu/alpha activity across gestures, as well as the relative lag between these events. To evaluate specificity to the mu/alpha frequency band, we also examined beta activity (15–24 Hz). This approach utilizes all available information (i.e., all inflection points across all gesture stimuli) pooled across participants, to test the generalizability of the temporal relationship across a range of gesture stimuli.

2. **Test whether inflection points evoke mu/alpha activity compared to gesture portions without inflection points** Within a subset of gestures, we extracted epochs containing an inflection point and compared them to epochs without an inflection point. Of interest was whether this contrast would show evoked mu/alpha activity at 11 Hz, the timing of evoked EEG activity, and the specificity of mu/alpha activity versus higher-frequency activity (e.g., beta). This approach tests whether EEG activity evoked by inflection point occurrence generalizes across participants.

---

**Material and methods**

The overview of the proposed experiment can be found in Fig. 1 and is described in more detail below. Using recorded videos from a freely-available gesture vocabulary, EEG data was acquired from participants while they passively watched the gesture movements. Motion trajectories for each gesture were coded for the timing of inflection points, and EEG data was processed to extract mu and alpha oscillations.
Participants

The data presented was collected from seventeen adult volunteers, as part of a larger ongoing project (Age: $M = 21$, $SD = 4$, range = 18–34; gender: 11 female, 6 male). All participants gave written consent, with ethical approval given by Purdue University Institutional Review Board. Preliminary analyses of four participants were published previously (Cabrera et al. 2017). Consecutive participants were run during an academic term as part of a larger series of studies, with no a priori power calculation. A sensitivity analysis ($\alpha = 0.05$, $1 - \beta = 0.80$, $N = 17$) indicated that with this sample size we would be able to detect linear correlations of at least $r = 0.48$ ($R^2 = 0.23$) and paired-sample comparisons of at least $d = 0.72$.

Gesture task

Participants passively viewed 20 videos taken from the ChaLearn 2013 stimulus set (Escalera et al. 2013). This data set is comprised by 20 gesture classes from Italian signs, which are communicative in nature (Fig. 2), even though understanding the meaning of the gesture was outside of the proposed task’s scope. Using McNeill’s classification of gestures (McNeill 1992), these 20 gesture classes can be considered a combination of iconic and metaphoric gestures; these hand gestures contain anthropological and cultural significance. There are approximately 14,000 gestures acquired from 30 different users in the public data set. For this particular experiment, only one instance of each gesture class was used, and they were separated into individual video files for a total of 20 unique videos. Each video was approximately 3 s in length and showed a single actor performing a gesture. Videos were presented in black and white and only the silhouette of the actor was visible, thereby making the physical motion of the gesture the most salient aspect of the video. Participants viewed each gesture a total of six times, and gesture order was randomized across participants. There was an inter-stimulus interval of 2 s between videos, during which a fixation cross was presented in the center of the screen.

Extracting motion information from gesture videos

The ChaLearn 2013 data set includes information regarding the position of 25 joints in the human body with respect to the sensor’s reference frame, additionally to the color and grayscale videos. With this information, it is possible to identify the location of the user’s hands in every frame of the selected gesture videos. The changes in position for both hands along the time axis (using all the frames in the video sequence) are used to describe the trajectory of the hands during the gesture execution.

By taking the derivative of both hands’ motion trajectory, it was possible to determine the occurrence of inflection points in motion. The inflection points are the local maxima or minima found in the derivative of the hands position over time. These are associated with abrupt changes in speed and orientation, which define the transition between concatenated movement phases (Kita et al. 1998). These abrupt changes that divide the gesture motion into phases are not dependent on the duration.
of the gesture (Viviani and Terzuolo 1982), but they are bounded by a sequential order (Bobick and Wilson 1997). The use of inflection points (IP) has been found in previous literature and associated with depicting the gesture’s intentions (Buchin et al. 2011; Despinoy et al. 2016; Loram et al. 2006).

Following this reasoning based on previous literature, inflection points are considered a salient characteristic or “placeholders” within each gesture class, under the assumption that the same placeholders would appear regularly across multiple examples of the same gesture class, and independently of variability associated with human nature. A different assumption tested, and concerning this work, is that such characteristics are highly correlated to the neural responses during the process of gesture perception.

The placeholders for each gesture class were obtained in the following way: the position of both hands was extracted for each frame of the video and the derivative of the motion was calculated. Then, the time occurrences for local maxima and minima on the derivative signal were extracted for the trajectories of both hands. For gestures that included motion in both arms simultaneously, the occurrence of inflection points were determined by merging the IP occurrences of each hand’s trajectory. By doing a transformation between the frame number where the inflection occurred, and the rate of frame per seconds (20 fps) of the videos, the occurrence of inflection points was determined for each gesture class.

The detailed information of the ChaLearn 2013 gesture sample used as stimuli for the experiment can be found in Appendix, including the time stamps of the extracted inflection points in the gestures’ motion trajectory. Overall, the extracted motion information can be summarized as follows:

- All gesture classes contained more than one inflection point. There were 3 gesture classes with only 2 IP, namely “G2”, “G19”, and “G20”.
- 17 out of the 20 gesture classes (85%) displayed at least 3 IP.
- Across all gesture classes, a total of 86 IPs were extracted and correlated with EEG data.

**EEG acquisition and analysis**

Continuous EEG responses were recorded using an ActiCap and the ActiCHamp amplifier system (Brain Products). The signal was digitized at 24-bit resolution and a sampling rate of 500 Hz. Recordings were taken from 32 scalp electrodes based on the 10/20 system, with a ground electrode at Fpz. BrainVision Analyzer (Brain Products) was used for offline analysis. Data were referenced to the average two mastoids (TP9/TP10) and bandpass filtered from 0.1 to 100 Hz. EEG signals were segmented from −1000 to 5000 ms relative to video onset, and ocular correction was performed using a regression-based algorithm. Artifacts were rejected within individual channels using an automated procedure: A 50 μV per ms voltage step with a maximal difference of 100 μV or more within 200 ms, and lowest activity of 0.5 μV within 100 ms. An average of 3.29% of data was rejected per participant at central and occipital sites (C3/Cz/C4, O1/Oz/O2).
Mu and alpha oscillations in response to each gesture were quantified using wavelet analysis. As opposed to a Fourier transform, which assumes a time invariant signal, wavelet transforms can be used to quantify changes in power over time. Briefly, a wavelet function at a specified frequency with a specified number of cycles was convolved with the EEG signal to quantify fluctuations in power at that frequency (Herrmann et al. 2005). Here, complex Morlet wavelets were applied separately to each gesture observation, using a frequency range of 1–20 Hz, frequency steps of 0.5 Hz, and Morlet parameter of $c = 7$. Segments were baseline corrected relative to the $−500$ to $−100$ ms window using a subtraction-based method (i.e., the average power at each wavelet layer from $−500$ to $−100$ ms was subtracted from each subsequent data point) and averaged separately for each gesture type (i.e., averaged across the 6 viewings of each gesture). Analyses focused on three layers to tease apart mu/alpha from higher-frequency activity (i.e., beta): (1) alpha was extracted as layer 21, with a central frequency of 11 Hz and a filter of 9.7–12.3 Hz, (2) low beta was extracted as layer 33, with a central frequency of 17 Hz and a filter from 15 to 19 Hz, and (3) middle beta was extracted as layer 41, with a central frequency of 21 Hz and a filter from 18.5 to 23.5 Hz. Each layer provided a time series indicating relative power at the central frequency at that time point and for that electrode.

Positive-going peaks occurring between 0 and 3000 ms were identified separately for each layer (mu/alpha, low beta, high beta) and each gesture stimulus, and the latencies of these peaks were used to characterize the time course of mu/alpha oscillations. Positive-going peaks were extracted from the waveforms using a Matlab® function called “findpeaks”; user-provided input parameters are used to extract local maxima from a vector as the function outputs, in terms of the peak’s absolute magnitude and the vector index of the occurrence. We focused on local maxima under the assumption that the gesture videos would elicit sustained decreases in mu/alpha activity, and that local maxima would reflect transient, oscillatory dynamics in the context of this sustained decrease in power. For our research, we kept the index information since it directly relates with time occurrence within the waveform. We extracted wavelet layers at electrodes C3/Cz/C4 (i.e., directly above the motor cortex), where mu activity is known to be maximal. For comparison, we also considered alpha activity at O1/Oz/O2 (i.e., an occipital cluster). We expected that these occipital electrodes would capture visual cortical activation associated with stimulus processing but not mu activity per se (Streltsova et al. 2010).

**Statistical analysis**

The overarching goal of these analyses was to test whether inflection points in gesture kinematics share a temporal relationship with peaks in mu/alpha activity that generalizes across gesture types. For this analysis, the three channels associated with the motor cortex responses (C3/Cz/C4) were averaged over and will be further referred to as Central Cluster; analogously, the channels associated with visual cortical activation (O1/Oz/O2) are further referred to as the Occipital Cluster. Inflection points extracted from one ChaLearn’s lexicon, are the local maxima found in the first derivative for the motion trajectory of the gesturer’s hands for each gesture class. Time stamps of inflection points from motion data and peak information from EEG data were plotted (time vs. time). Trend lines and coefficients of determination were obtained to check correlation on the plotted data. Additional linear regression models were fitted, and the models’ coefficients were determined to check the significance of their predictive nature. The aforementioned metrics were used in three different analyses for all three layers discussed previously (alpha, low beta, and middle beta):

**Standard analysis**

For this analysis, the underlying assumption is that salient occurrences in the gesture motion trajectory, hypothesized to be the extracted inflection points, generate a subsequent peak in neural response as oscillations in the EEG data. The regression models were fitted to determine the timing of peaks in EEG data using the timing of inflection point occurrence as the predictors.

**Reverse analysis**

This analysis considers the opposite assumption: peaks in EEG data precede occurrences of inflection points in gesture motion. Given the nature of the standard analysis, this analysis performs a different measurement with the intention of finding an alternative significance of the proposed scenario. If this analysis shows no meaningful correlation, it would discourage the possibility of assigning the correlations found in the standard analysis as spurious relationships. The regression models were fitted to determine the occurrence of inflection points using EEG peak data as the predictors.

**Random analysis**

For this analysis, random EEG data was considered under the same assumption as the standard analysis, occurrences in gesture motion precede EEG peaks. To achieve this, the same EEG acquired data was flipped in time and motion
information of gesture class $i$ was paired with EEG data from gesture class $j$, where $i \neq j$. By shuffling EEG data, random oscillations were achieved while maintaining realistic variability within the data. The purpose of this manipulation of the EEG data, is to determine whether a relationship would arise from the analysis of inflection points used as predictors of EEG oscillations, if the EEG data was random, without generating “synthetic” EEG data.

**Multi-level hierarchical regression**

This analysis is proposed as an alternative check of the influence of inflection point occurrences in motion as predictors for EEG peaks, using the data from the standard and reverse analysis. The hierarchical regression allows to modify the predictors to generate a model and check the significance of the correlation and the coefficients within the models. The steps in the multi-level regression were as follows: first, only information from the reverse analysis in motion was used to predict EEG peaks, then a model was fitted including the motion information from both the standard and reverse analysis. The SPSS software was used for the statistical analysis, and residuals and normality assumptions were checked.

**Alternative analysis**

The analyses described above focus on EEG epochs segmented based on the onset of the gesture stimuli, mapping the occurrence of inflection points onto subsequent mu oscillations in the observed EEG waveform. While these analyses are able to capture the dynamics of mu oscillations throughout the presentation of each gesture stimulus, they do not directly address whether oscillatory mu activity was evoked by the occurrence of inflection points per se, or whether these effects generalize across individuals. To test this latter question, we conducted a separate set of analyses using EEG epochs segmented based on the occurrence of inflection points. We focused on gesture stimuli containing (a) a 1300 ms epoch surrounding a single inflection point (− 500 to 800 ms relative to the inflection point) and (b) a separate, non-overlapping 1300 ms epoch with no inflection point. Four gesture stimuli met these criteria: G1, G8, G11, and G19. Following segmentation, the identical wavelet transform as described above was applied. Epochs with and without inflection points were averaged separately for each participant across the four gesture stimuli, and evoked EEG power was evaluated by calculating the contrast across these matched epochs with a paired-sample $t$ test, calculated separately for each frequency band (1–24 Hz, in steps of 0.5 Hz) and time point (0–800 ms, in steps of 2 ms). To correct for multiple comparisons, we controlled the false discovery rate across frequency steps ($q < 0.05$) (Benjamini and Hochberg 1995) set a minimum time duration of 20 ms (i.e., 10 consecutive contrast across time points). Contrasts were calculated using Brainstorm (Tadel et al. 2011, 2019).

**Descriptive analysis of captured EEG responses**

The acquired EEG responses were averaged across participants for each gesture class to test if a pattern can be extracted from commonalities in mu activity. The occipital cluster showed a decrease in EEG power at the alpha layer (Fig. 3, bottom). This is consistent with corresponding alpha activity and signals visual cortex activation associated with viewing a stimulus. This response was sustained, averaging over all gesture classes, within the 400–2500 ms time range, indicating continuous visual activity for the period of time in which participants attended to the gesture videos and highlighted with a gray box in Fig. 3.

Regarding the central cluster, a different pattern was observed. There was evidence of mu suppression along the 11 Hz band between 200 and 2500 ms, within the gestures’ duration. The most accentuated cluster of mu suppression peaked at approximately 400 ms, highlighted in Fig. 3 (top) by a gray box. This indicates an automatic suppression of mu activity, independent of any salient characteristic of each gesture class. Expanding this observation, 13 out of the 20 gesture classes (65%) showed a mu peak even before any inflection points were registered in the gesture class with an average response of 280 ms. This is consistent with previous results found in a preliminary study reported by Cabrera et al. (2017). Additionally, a previous EEG study also shows mu suppression within the first second of gesture viewing, regardless whether the gesture was communicative or meaningless in nature; this is a general response within the motor cortex distinct from visual cortical response apparent at occipital electrodes (Streltsova et al. 2010).

Breaking down the data per each gesture class, 95 peaks for mu activity were detected, defined as local maxima occurring after stimulus onset (counted after the first IP occurred in each gesture class). These positive-going peaks of mu activity represent the first evidence of dynamic mu oscillatory activity associated with gesture interpretation. All gesture classes elicited a first peak in mu activity with an associated average latency of 420 ms ($SD = 150$ ms) between IP and mu peak occurrence.

For all 86 IP occurrences considered in the stimulus set, there was a subsequent mu peak with an average latency of 491 ms ($SD = 213$ ms) relative to the IP occurrence. There were 12 peaks in mu activity that did not correspond directly with an IP occurrence: 9 of them occurred after the last IP had occurred, while 3 (3.2% of all mu peaks) were elicited in between IP, within the time interval the gesture was being executed (these were found in only 3 of the gesture classes). The average latency for these non-corresponded peaks was 867 ms. Given the nature of the
gesture vocabulary, each gesture began and ended with the performer doing the same resting pose: arms extended with the hands at hip level. The occurrence of a late mu peak (45% of the gesture classes) could be attributed to recognizing the resting pose at the end of the executed gesture.

Two gesture examples are examined in more detail, namely G7 and G12. Figure 4 shows the time course of mu activity for each gesture, using the 11 Hz wavelet. Frames from the gesture video corresponding with the IP found showed above each graph are mapped to the signal time series with indicative arrows. The corresponding mu peaks are labelled with a red circle. Note that positive power is plotted down in both mu graphs. The latencies were 426 ms for G7 and 348 ms for G12, respectively.

These two gestures were used to compare against previous preliminary results reported (Cabrera et al. 2017). These two gestures also showed correspondence previously between the number of inflection points and the observed peaks in mu activity. These results continue to be consistent with the proposition that inflection points function as placeholders within gesture expression, and that these placeholders modulate activity within the motor cortex that is associated with gesture processing and interpretation.

Results

Statistical analysis

Standard analysis

For this analysis, the time occurrence of an inflection point in motion trajectory was used for the abscissa axis while the time occurrence of a peak in EEG layer, either in the averaged central cluster corresponding to activation in the motor cortex in the alpha layer (Fig. 5, left) or the averaged occipital cluster corresponding to activation in visual cortex for the alpha layer (Fig. 5, right). Lines were fitted and the coefficients of determination $R^2$ were obtained for this scenario, $R^2 = 0.9215$ and $R^2 = 0.8592$, respectively ($p < 0.01$) per channel type.

A high correlation was found between the incidences of salient motion components with peaks in neural oscillations.
Another result worth highlighting from these fitted lines are the independent terms obtained, associated with time lags or reaction time between the motion stimuli and the EEG response. In the case of the occipital cluster the lag was found at $t_O = 465.01\text{ms}$, while the lag for the central cluster was found at $t_C = 535.93\text{ms}$. An additional analysis was conducted, acknowledging the fact that several IP with EEG spikes are nested within the same gesture. By considering only the occurrence of the first inflection point and the first EEG peak for the central cluster, a similar behavior is observed with the regression model showing a lag of $t^{*}_C = 428.96\text{ms}$ with $R^2 = 0.8696$ ($p < 0.05$).

The standard analysis was also conducted for the other two extracted wavelet frequency layers: low beta (Fig. 6) and middle beta (Fig. 7); similarly, regression models were fitted using the occurrence of inflection points in motion as predictors for peaks occurrences in EEG waveforms. Low beta fitted lines and $R^2$ determination coefficients are shown in Figs. 6 and 7: the left scatterplot corresponding to the central cluster and the right to the occipital cluster. Regression models showed no significance for either cluster in the low beta (central, $p = 0.132$ or occipital, $p = 0.247$) or middle beta (central, $p = 0.216$ or occipital, $p = 0.309$). On the low beta
layer, the determination coefficients were $R^2 = 0.706$ for the central cluster and $R^2 = 0.682$ for the occipital cluster. On the middle beta layer, the determination coefficients were $R^2 = 0.666$ for the central cluster and $R^2 = 0.7306$ for the occipital cluster.

These results indicate that higher correlation between oscillations in mu activity and salient motion features can be found primarily within the alpha wavelet layer (11 Hz).

**Reverse analysis**

For this analysis the time occurrences of inflection points in gesture trajectories were plotted as a function of EEG peaks occurring in the averaged central cluster, as shown in Fig. 8. The fitted trend line and $R^2$ coefficient obtained from this analysis show no correlation ($R^2 = 0.006$). No statistical significance was found for the central cluster at the alpha layer as predictors for inflection points in motion in this instance ($p = 0.8157$). This analysis contests the idea of correlation simply due to relating contiguous measurements of IP against EEG, given the false reciprocal. The absence of correlation provides a basis to reject the assumption that EEG oscillations precede the occurrences of inflection points in gesture trajectories.
Random analysis

In this analysis, the plotting scheme was the same as the standard analysis, showing central cluster peak occurrences as a function of inflection points in motion in Fig. 9. The fitted trend line and $R^2$ coefficient obtained from this analysis show no correlation ($R^2 = 0.0859$, $p = 0.7193$). The random nature in the EEG data used for this analysis showing no correlation with salient points of the gesture trajectories, provides strength to the previous results found with the standard analysis discarding a spurious patterning.

Multi-level hierarchical analysis

The conducted analysis evaluated the change in $R^2$ for models predicting the occurrences of peaks in EEG central cluster ($\text{EEGC}_c$) at the Alpha layer using the occurrences of inflection points from the standard analysis ($\text{Motion}_{\text{rev}}$) and the reverse analysis ($\text{Motion}_{\text{reg}}$). Normality assumption was checked using Q-plot and residuals. Pearson correlation was found at 0.94 for $\text{Motion}_{\text{reg}}$ and 0.094 for $\text{Motion}_{\text{rev}}$. Additional results related to the fitted models and their significance are shown in Table 1.

Alternative analysis: evoked mu/alpha

This analysis focused on the within gesture contrast of matched epochs with and without an inflection point. The time–frequency decomposition is presented in Fig. 10. At both the central and occipital clusters, there were evoked EEG responses following the occurrence of an inflection, as compared to matched epochs from the same gesture stimuli containing no inflection points. These responses were statistically significant ($p < 0.05$) after correcting for false discovery rate across frequency band. At the central electrode cluster, the evoked responses showed relative temporal and spectral specificity: there was an increase in mu activity from 308–364 ms spanning 8.5–13.5 Hz, and an increase in beta activity in the same time range spanning 18–23.5 Hz (Fig. 10, top). At the occipital cluster there was broader, sustained activity, including: an increase in alpha activity from 226 to 800 ms spanning 9.5–12.5 Hz, and an overlapping increase in beta activity from 306 to 800 ms and from 13 to 24 Hz (Fig. 10, bottom).

Discussion

The results presented in this work are consistent with previous results reported in literature, showing that mu rhythms are sensitive to gesture observation (Braadbaart et al. 2013; Quandt et al. 2012). While previous studies have focused on the average reduction of EEG power, particularly mu suppression, while the stimulus is presented, this work focused on the oscillatory nature of mu activity and determining whether the occurrence of peaks in mu power would correspond to salient characteristics found in the gesture execution. The mentioned gesture characteristics are the points in the motion trajectory that display abrupt changes in either magnitude or orientation, namely the inflection points. These abrupt changes are considered relevant, connecting consecutive gesture phases (Kita et al. 1998). A linear relationship was found between the times when the gesture executed exhibited an inflection point and the time occurrences of peaks in EEG signals, particularly from the central and occipital electrodes. Average lags of 465 ms for the occipital electrodes and 535 ms for the central electrodes were found, and effects were specific to the mu/alpha band (11 Hz) and not beta. These effects show a temporal association between gesture kinematics and mu/alpha oscillatory activity that generalizes across a wide range of naturalistic gesture stimuli, but not necessary across participants. To
test the latter, we also examined EEG activity evoked by inflection points. When segmented based on inflection point occurrence rather than gesture stimulus onset, two evoked responses were observed at the central cluster: increased mu activity and beta activity, both from 308 to 364 ms. Broader and more sustained evoked activity was observed at the occipital cluster, including an increase in alpha activity beginning at 226 ms. These results show a close correspondence between gesture motion and EEG oscillations across the time of respective gesture execution and perception.

The detected oscillatory patterns mention both central and occipital electrodes. One main difference between the Streltsova study (2010) and ours was they considered average EEG power in fixed 1000 ms intervals, regardless of the specific motion trajectories of the gestures themselves, while we used wavelet analysis to retain temporal resolution and detect idiosyncratic peaks within the EEG waveform for each gesture stimulus. Close responses were found between the timing of peaks at central and occipital electrodes with the timing of inflection points in gesture motion trajectories. This could be interpreted as coordinated activation of visual and motor cortices involved in gesture processing. Given the earlier timing of the peaks found in occipital electrodes to the ones found in central electrodes (465 vs. 535 ms for the peak latency analyses; 226 vs. 308 for the analyses of evoked activity), this suggests visual cortical activity was driving motor cortical activity with a lag of approximately 80 ms, although this needs to be verified in future research. We note here that the peak analyses focused on local maxima in the waveforms (i.e., transient increases in mu/alpha activity occurring within the broader context of ongoing mu suppression), but by definition these oscillations are preceded and followed by local minima, which would yield different estimates of peak timing.

A descriptive analysis of the results found further describes the close correspondences between inflection points in motions with peaks in EEG signals associated to motor and visual cortices, focusing more on the differences among gesture classes. All inflection points extracted from gesture motion were matched to central cluster oscillations in the alpha layer. There were additional oscillations in the central cluster that did not correspond to a preceding inflection point and occurred before the next in three of the gesture classes. Only 3.2% of all mu peaks found averaging over all 17 participants, displayed such behavior with an average lag of 867 ms. These oscillations in the central cluster may be attributed to other salient characteristics of the gestures that were not considered in this work: hand shape configuration or relative position to the gesturer’s body, among other features relevant in sign language. In regards to the peaks found
during the gesture, in between IP, there could be other salient characteristics in those three gesture classes that were not related to the gestures’ trajectories, for instance an additional response related to attention or learning which also manifest in the same band (Başar 2010). Additionally, nine gesture classes exhibited a mu peak towards the end of the gesture video, when the gesture was returning to a resting pose after executing each gesture class, which did not correspond to any extracted inflection points in motion. One interpretation of this could be related to recognizing the resting pose itself.

The conducted analysis paves the way towards future on gesture recognition. The EEG oscillations found related to salient points in motion can be used to build a global model of a compact representation of a gesture. This “mental snapshot” of the gesture includes rich information about the gestures’ modes of variation. By “unfolding” this representation, human-like perception and movement production could be transferred to machines, making human–machine interaction more intuitive and natural. Understanding how people learn to mimic and produce gestures and transferring those techniques to machines can be considered a multidisciplinary effort, requiring expertise from psychology, motor control, human–machine interaction, machine learning and linguistics. The current findings are limited, among other things, by the salient points considered to describe the gesture’s model; the use of motion trajectories can be used for gestures involving motion in the upper limbs, however the configuration of the hand or changes in the position of the fingers are overlooked presently. There is another limitation related to the communicative nature of the gestures. Given that no comparison has been conducted yet on meaningless gestures, it would be beneficial to test if the visual–motor EEG oscillations are specific to communicative gestures or generalize to other gesture typologies (e.g. non-communicative gestures), and furthermore to all instances of biological motion, regardless of communicative intent.

The current findings are consistent with the premise that inflection points are salient characteristics involved in gesture perception, recognition and repetition. As such, the oscillation found in EEG mu rhythms are sensitive to occurrences of inflection points during gesture processing. The cognitive processes related with gesture production and perception could potentially be considered as a source of features for gesture recognition.

Conclusions

The purpose of this paper is to validate the claim that an executed gesture contains a finite set of salient points that produce neural signatures, involving oscillations in EEG signals, associated with cognitive processes. Those signatures may be used to keep in memory intrinsic characteristics of the gesture. This was achieved by determining whether a relationship exists between the time occurrences of salient events in both motion trajectories of performed gestures, and mu oscillatory rhythms for participants observing the gestures.

A statistical analysis was conducted across gestures using linear regression and three different conditions. Peaks in the EEG signal at central electrodes and occipital electrodes were used to isolate the salient events within each gesture. Average lags in EEG power oscillations were detected at 465 ms and 535 ms after inflection points at occipital and central electrodes, respectively. Separate analyses conducted across participants indicated that inflection point occurrences evoked occipital alpha activity at 226 ms and central mu activity at 308 ms. The results suggest that coordinated activity in the visual and motor cortices are highly correlated with key motion components within gesture trajectories, and it is consistent with the proposal that inflection points operate as placeholders in gesture recognition. The potential of this work is that it provides evidence that inflection points are key points in gesture trajectories, which can encapsulate the gestures in human cognition. Therefore, these points can be used to capture large variability in each gesture while keeping the main traits of the gesture class.

Acknowledgements

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs under Award No. W81XWH-14-1-0042 (JPW is Principal Investigator). Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense.

References

Arbib MA (2005) Interweaving protosign and protospeech: Further developments beyond the mirror. Interact Stud 6:145–171
Avenanti A, Bueti D, Galati G, Aiglotti SM (2005) Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. Nat Neurosci 8(7):955
Başar E (2010) Brain–body–mind in the nebulous Cartesian system: a holistic approach by oscillations. Springer, New York
Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57(1):289–300
Bobick AF, Wilson AD (1997) A state-based approach to the representation and recognition of gesture. Pattern Anal Mach Intell IEEE Trans 19:1325–1337
Braudbaart L, Williams JH, Waiter GD (2013) Do mirror neuron areas mediate mu rhythm suppression during imitation and action observation? Int J Psychophysiol 89:99–105
Buchin K, Buchin M, Gudmundsson J, Löffler M, Luo J (2011) Detecting commuting patterns by clustering subtrajectories. Int J Comput Geom Appl 21:253–282
Cabrera ME, Novak K, Foti D, Voyles R, Wachs JP (2017) What makes a gesture a gesture? Neural signatures involved in gesture recognition. In: 2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017), pp 748–753
