

Sensory-Motor Gestalt: Sensation and Action as the Foundations of Identity, Agency, and Self

Jasmine A. Berry and Francisco J. Valero-Cuevas

Posted Online July 14, 2020
https://doi.org/10.1162/isal_a_00340

© 2020 Massachusetts Institute of Technology Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license

ALIFE 2020: The 2020 Conference on Artificial Life
July 2020
p.130-138

Download

Favorite

Download Citation



Conference Date: July 13 - 18, 2020
Location: Online
Date Published: July 2020

Editors: Josh Bongard, Juniper Lovato, Laurent Hebert-Dufrésne, Radhakrishna Dasari and Lisa Soros

More About ALIFE 2020 ▾

[The ALIFE 2020 website](#)

More About ISAL ▾

[The International Society for Artificial Life website](#)

Abstract Authors

Body movement and proprioception are inextricably linked. Movement produces continuous high-dimensional ensembles of afferent information that provide an *internal proprioceptive body representation* and its relationship to the environment. Motor function is amenable to recording and interpretation and has been relatively well studied. However, we do not yet understand how physiological proprioceptive afferents contribute to internal body representations, neuromuscular control, and even a sense of agency and self. Proprioceptive and motor signals have often been seen as separate, and to be combined mainly to close feedback loops for neuromuscular control. In contrast, 'active sensing,' is an emergent concept for dynamically blending sensory and motor signals. We extend and formalize active sensing into an integrative approach — born out of a neuromechanical perspective — that sees proprioceptive and motor signals as integral parts of the same functional and perceptual continuum we call the *Sensory-Motor Gestalt*. The Sensory-Motor Gestalt combines formalisms of physics, state estimation, biomechanics, differential geometry, and physiology to understand the emergence of the self in the context of proprioception and motor actions in the physical world. Proprioception, by defining *body state*, defines feasible (continuous or discrete) motor actions compatible with that state and the environment. Conversely, motor actions produce subsequent, often predictable, body states. This syntactical relationship leads to an epistemological continuum that spans body state, feasible behavior, agency, identity, and sense of self in organisms and robots.

Sensory-Motor Gestalt: Sensation and Action as the Foundations of Identity, Agency, and Self

Jasmine A. Berry¹ and Francisco J. Valero-Cuevas^{1,2,3}

¹Department of Computer Science

²Department of Biomedical Engineering, ³Division of Biokinesiology and Physical Therapy
University of Southern California
jasminab@usc.edu

Abstract

Body movement and proprioception are inextricably linked. Movement produces continuous high-dimensional ensembles of afferent information that provide an *internal proprioceptive body representation* and its relationship to the environment. Motor function is amenable to recording and interpretation and has been relatively well studied. However, we do not yet understand how physiological proprioceptive afferents contribute to internal body representations, neuromuscular control, and even a sense of agency and self. Proprioceptive and motor signals have often been seen as separate, and to be combined mainly to close feedback loops for neuromuscular control. In contrast, ‘active sensing,’ is an emergent concept for dynamically blending sensory and motor signals. We extend and formalize active sensing into an integrative approach—born out of a neuromechanical perspective—that sees proprioceptive and motor signals as integral parts of the same functional and perceptual continuum we call the *Sensory-Motor Gestalt*. The Sensory-Motor Gestalt combines formalisms of physics, state estimation, biomechanics, differential geometry, and physiology to understand the emergence of the self in the context of proprioception and motor actions in the physical world. Proprioception, by defining *body state*, defines feasible (continuous or discrete) motor actions compatible with that state and the environment. Conversely, motor actions produce subsequent, often predictable, body states. This syntactical relationship leads to an epistemological continuum that spans body state, feasible behavior, agency, identity, and sense of self in organisms and robots.

Introduction of Sensory-Motor Gestalt: Origin and Definition

“To understand is to perceive patterns.” -Isaiah Berlin

Our computational model of the self begins with *Gestalt Theory*. Gestalt (pronounced gə-’shtält), a concept originating in Austria and Germany, roughly translates to shape, form, configuration, and unified whole. XX-century German psychologist Max Wertheimer utilized this definition to originally present the Gestalt laws (or principles) of grouping for pictorially detailing how the human eye perceives visual elements (Wertheimer, 1923). These laws are fundamental rules illustrating how humans recognize elements and objects in their visual scene as organized patterns with

meaning. The Gestalt theory of the mind and brain intends to form an understanding of how humans and animals 1) comprehend what they are perceiving and 2) obtain meaning from the world with disordered visual stimuli.

In its original formulation, Gestalt laws aim to reduce complex visual scenes into simpler, less complex shapes to convey an image’s meaning in a single formation instead of disparate smaller elements. Hence by being a critical aspect of the perception of patterns into a coherent whole for context and meaning, Gestalt plays an important role in combining epistemology (i.e., study of knowledge and how does one come to “know”) and ontology (i.e., study of what is the nature of the self) (Guarino et al., 2009). This paper proposes underlying mechanisms for brain-body dynamics to merge proprioceptive and motor elements into an epistemological continuum from sensory and proprioceptive input, to state of the body, to feasible motor action, to useful behavior, to the sensory consequences of action—and then on to more abstract notions of agency, identity and sense of self in organisms and robots.

In this study, we focus on proprioception as spike trains from muscle spindles (II and Ia) and Golgi tendon organs (GTOs). They primarily encode muscle fiber length and velocity, and tendon tension, respectively. These proprioceptive signals are known to inform various perceptual modalities of body state (e.g., postures, movements, forces, limb stiffness, alertness). Recent exponential growth in literature relating proprioception to subjective experiences (Figure 1) may suggest that the debate about the emergence of the self is advancing. We hypothesize that Gestalt laws can be applied to organizing these physiologically-tenable proprioceptive signals to construct a *totality* of what is perceived as the active body. We seek to do the same for motor actions by developing a mathematical description of the set of plausible motor actions conditional on proprioceptive signals.

Our prior simulation work characterized the high-dimensional, non-linear, time-varying manifolds of muscle spindle afferents (Ia and II encoding, roughly, muscle fiber contraction velocities and lengths, respectively) that emerge during movement of a planar multi-muscle

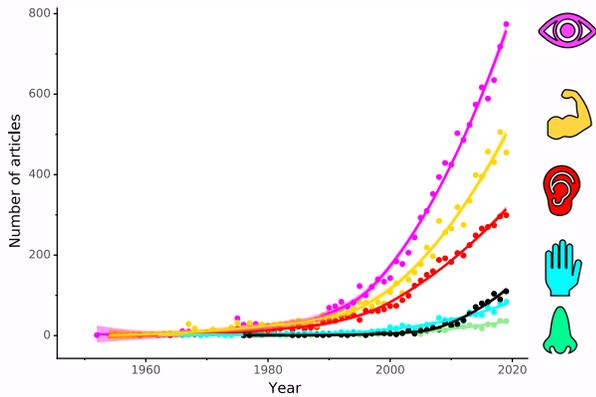


Figure 1: Published article count per year that indicate association between subjective experiences and sensory modalities. This chart similarly models a previous search conducted by Faivre et al. (2017), in which the number of articles published mentioned the words “awareness” or “consciousness” in conjunction with words denoting each sensory modality: “Visual” or “vision” (magenta), “proprioception” or “movement” (yellow), “auditory” or “audition” (red), “touch” or “tactile” (cyan), “olfaction” or “olfactory”, and “multisensory” (black). Our PubMed search extended the year range from 1950-2019 and added proprioception. Along with vision, proprioception showed a significant increase in documented work.

limb (Berry et al., 2017). We provided initial evidence that high-dimensional muscle spindle proprioception defines limb states that reflect the consequences of motor actions. We now extend that work by emphasizing that each limb state, by its physical nature, only has a well-defined set of feasible motor actions. This results in a formal sensory-to-motor-to-sensory map that defines both the current proprioceptive states and feasible motor commands (i.e., plausible motor actions) that will lead to new (but expected) proprioceptive signals. The Sensory-Motor Gestalt applies to both biological and engineered agents where the concepts of state, observability, and controllability are intimately related; therefore, providing a basis for constructing an artificial core of state, agency, identity and ultimately self.

To our knowledge, this is a first attempt to formally apply the laws of Gestalt to the encoding of the sense of agency, identity and self via proprioception. This article first builds the concept of Sensory-Motor Gestalt from the generic Gestalt theory. We then interpret the Sensory-Motor Gestalt in the form of mathematical encoding for each of the core laws, which may be integrated to form the sensorimotor self. How the self is directly related to sensorimotor experiences of neuromuscular systems is then explored while providing sample platform applications to support Sensory-Motor Gestalt, both biological and in bioinspired robots.

Sensory-Motor Gestalt: Applying Gestalt Laws to Sensorimotor Function

“The whole is greater than the sum of its parts” is the popular adage Gestalt psychology is best known for. It emphasizes the fact that although a sensory experience can be disassembled into individual components (i.e., stimuli), the way in which those components coalesce together generates properties and qualities of the whole that only exist independently of their components. Stimuli patterns presented as a whole often prompts a more meaningful perceptual response. As alluded to earlier, Gestalt theory is typically associated with the visual sense and visual perception (e.g., object and shape recognition, coloring, arrangement of parts) that is used to process graphic designs and images (Figure 2). Rarely has Gestalt theory been applied to other sensory modalities such as haptic (Chang, 2007), auditory, and olfactory senses which can all be topographically represented on a multi-dimensional space in the depiction of manifolds (i.e., coherent and continuous lower-dimensional subspaces embedded in a higher-dimensional space). Interestingly, the functional mechanisms of Gestalt laws are active in other cortical areas of the brain and not solely in the visual processing centers. As the brain’s neural processing is responsible for stitching together the visual scene of the external world in the primary sensory cortices and also seamlessly binding raw multisensory information to project a single unified experience, we theorize there are benefits in extrapolating ideas of Gestalt laws of perceptual grouping from vision to other modalities of the body such as the somatosensory system. Gestalt theory typically consists of five core laws that govern the fundamental organization of perception: Laws of Proximity, Similarity, Closure, Continuity, and Prägnanz.

Law of Proximity

Within visual perception, objects in space or points on a plane that are near or proximate to each other have a tendency to be grouped together in a single unified set. Conversely, points that are further apart have a lesser likelihood to be viewed as conjoined (see Figure 2a). This law is useful for organizing information with increased speed and efficiency. There are several ways in which proprioceptive information can also be processed to yield proximity metrics. One of the earlier attempts to address the Law of Proximity is the Pure Distance model (Kubovy et al., 1998), which attempts to quantify visual proximity grouping in dot lattices with an attraction function that measures the probability distribution of grouping.

Several algorithms can process proprioceptive stimuli in this manner. Consider our prior work (Berry et al., 2017) on the simple case of spindle model output of a single muscle fiber, which is 2-D Ia and II afferent spike trains over time. Figure 3 shows a higher-dimensional case of a simulated human arm. When examining proprioceptive signals

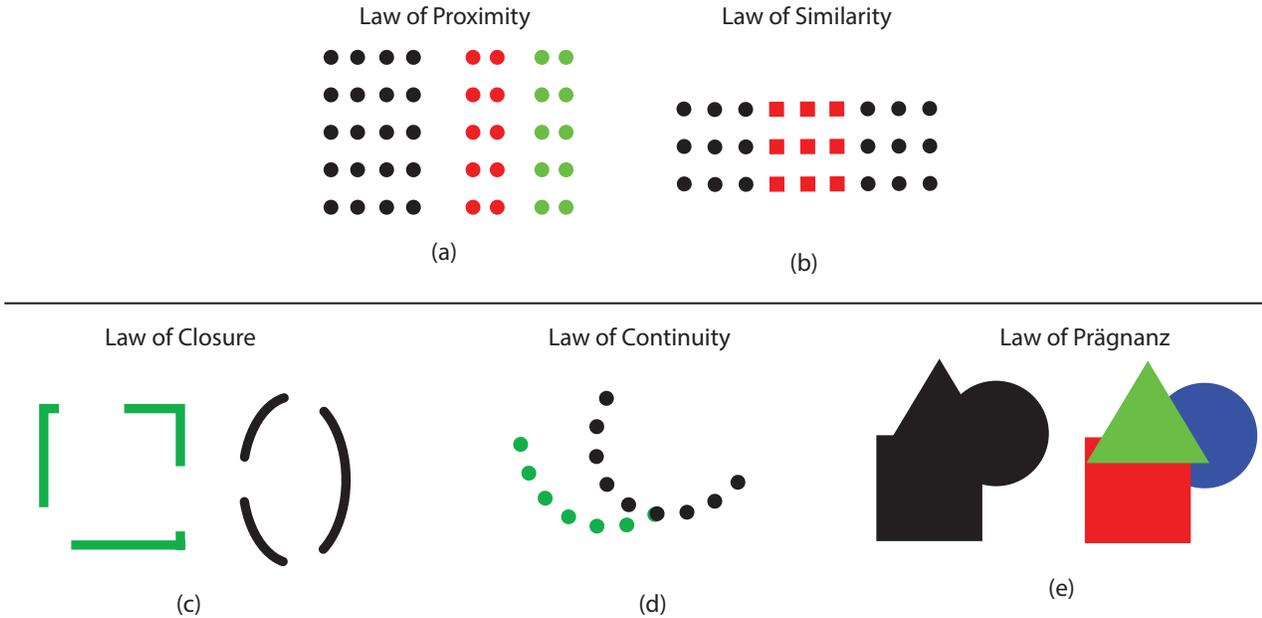


Figure 2: Gestalt laws of perceptual organization for topological manifold data can be applied using these core laws: **(a) Law of Proximity** aims to group elements together based on spatial closeness. Black dots, red dots, and green dots are perceived as separate groups due to the nearness of columns. **(b) Law of Similarity** groups the elements of black dots and red squares as separate sets, although spatial distance between each element is consistent. Shape, orientation, and color are the distinguishing factors. **(c) Law of Closure** prompts pattern perception of a green square and black oval despite the non-continuous outline and presence of gaps. **(d) Law of Continuity** perceives the figure as a green dotted line and a separate black dotted line due to the observed fluid connection of continuity and direction. **(e) Law of Prägnanz** (i.e., pithiness, conciseness, or Good Form) takes the abstract shape, as depicted on the left, and perceptually reorganizes them into a simple, more recognizable forms as is depicted on the right with the colored circle, triangle and square.

that are encoded as spike trains in units of pulses per second (pps), we are presented with unlabeled sample points (x_1, x_2, \dots, x_n) , where n is the set of observations, that can be further mathematically expressed to form representations. Since the notion of proximity is to associate observed points by measurement of Euclidean distance, then a standard unsupervised algorithm such as the K-Means clustering (i.e., a simplified version of vector quantization) proves to be sufficient for revealing underlying data structure.

Law of Similarity

Elements (e.g., points) that are similar in visual appearance in at least one degree with alike components are more likely to be grouped and organized together perceptually. The Law of Similarity generally spans the attributes of orientation, texture, color, and shape (see Figure 2b). There are ways to apply this law to the manifolds produced by proprioceptive signals. Considering the contours and curves that emerge from the collection of proprioceptive manifolds (e.g., Figure 3), shape is the most applicable attribute when measuring for similarity. Shape dimensions, such as curvature and elongation, can be perceived as integral dimensions and also used for comparison for similarity. In a similar fash-

ion that Boyer et al. (2011) quantifies geometric similarity of anatomical surfaces and morphological identification, we can apply statistical analysis when viewing the Ia and II stimuli as a collection of discrete or continuous points on an anatomical surface. Measures of similarity of each afferent signal across various tasks can be applied across the collected time-series data using signal processing. Comparable to the Law of Proximity, K-means clustering may also be used here if measuring 'similarity' of clusters by its relation to Euclidean distance of data points.

$$\cos \theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|} = \frac{\sum_1^n x_i y_i}{\sqrt{\sum_1^n x_i^2} \sqrt{\sum_1^n y_i^2}} \quad (1)$$

In our example, let x and y be two vectors of afferent spike trains, Ia and/or II. The cosine similarity function is a measure of similarity that can be used to compare afferent signals in the inner product space. Using the cosine measure, we have Eq. (1) where $\vec{x} \cdot \vec{y} = \sum_1^n x_i y_i = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$ is the dot product of the two vectors. A cosine similarity, $\cos \theta$, value closer to 1 indicates a higher propensity for perceptual clustering along the manifold. The convolution function would be another choice that quantifies similarity over time for all possible lags between

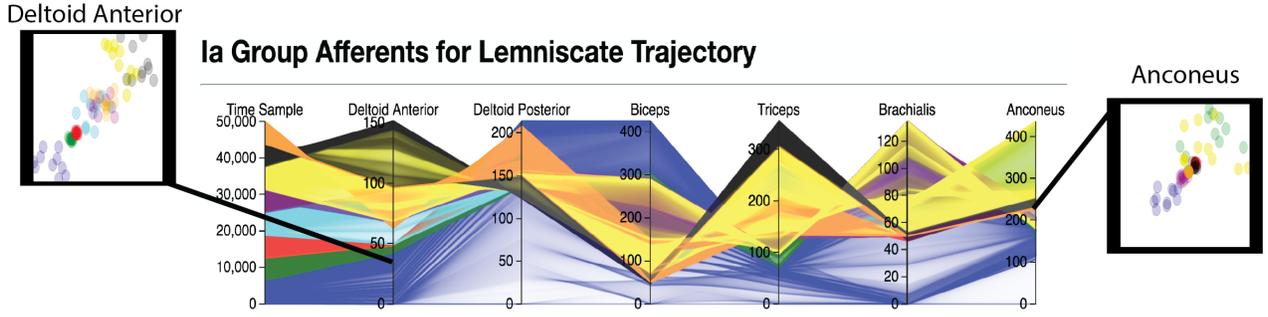


Figure 3: Spike trains from spindle afferents produce an evolutionary high-dimensional time-varying manifold of raw afferent information that is distinct for different arm movements. Using parallel coordinates, we show the Ia Group Afferent in 50,000 time samples for the case for a 6-muscle, 2-joint simulated planar arm performing the Lemniscate (figure-of-eight) trajectory with the end point (Berry et al., 2017). The coordinates are colored according to the segmented locations within the duration of the Lemniscate trajectory. The shadow boxes to the left and right of the manifolds are scaled-down sample snapshots of the data for the Deltoid Anterior and Anconeus muscles, respectively; ultimately revealing their specific cluster ranges.

signals.

Another option is cross correlation. It compares the time-series of afferent data across tasks, and is represented as the ratio in Eq. (2), where n is the total number of data point indices recorded per task cycle. Both x_i and y_i are the individual spindle afferent sets, Ia and II, respectively. A temporal shift delay, phase lag τ , of the output cross correlation, R_{xy} , measure is applied to determine where the correlation of the data is maximized, as shown in Eq. (3).

$$R_{xy}(\tau) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

$$\tau_{estimated} = \arg \max_{\tau \in \mathbb{R}} (R_{xy}(\tau)) \quad (3)$$

Magnitude-squared coherence is similar to correlation except that signals are compared in frequency ω , instead of time space, as shown in Eq. (4), which values satisfy $0 \leq C_{xy}(\omega) \leq 1$. $S_{xy}(\omega)$ represents the cross-spectral density between x and y , while $S_x(\omega)$ and $S_y(\omega)$ are the autospectral densities for their respective signals.

$$C_{xy}(\omega) \triangleq \frac{\|S_{xy}(\omega)\|^2}{S_x(\omega)S_y(\omega)} \quad (4)$$

$$\begin{aligned} S_{xy}(\omega) &= \int_{-\infty}^{\infty} R_{xy}(t) e^{-j\omega t} dt \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x(\tau) \cdot y(\tau + t) d\tau \right] e^{-j\omega t} dt \end{aligned} \quad (5)$$

Lastly, Kullback–Leibler (K-L) divergence is a means to quantify the likelihood that the statistics of a given process are similar to that of another, Eq. (6). Probability distributions P and Q are measured in comparison to reveal the relative entropy. This is particularly useful because it measures

how much information is lost when we approximate distributions.

$$D(P \parallel Q) = \sum_{x \in X} P(x) \log \left(\frac{P(x)}{Q(x)} \right) \quad (6)$$

Law of Closure

The Law of Closure is the tendency to complete unfinished or partially obscured objects. Here, incomplete figures are seen as complete or whole as depicted in Figure 2c. Warshall’s Algorithm (Warshall, 1962) may address this through its approach in computing the transitive closure of a node relation in a graph. We can envision, that as clusters are being formed via other laws, state nodes will eventually emerge from the aggregate data. To establish state transitions from one afferent cluster to another, the Warshall algorithm can determine whether a vertex j is ‘reachable’ from another vertex i for all vertex pairs within the graph. This measure of reachability will serve as the transitive closure, indicating directions and where paths exist for point-to-point movement across the manifolds.

This law states that, given available information, there is the expectation (based on prior personal experience) of closure when a fragmented version is presented. Bayes’ Rule is a formal way to represent such expectation in the case of visual information, visuomotor perception (Körding and Wolpert, 2004), and now proprioception. Bayes’ rule states that we can obtain the posterior distribution (the probability of a given body state given current proprioceptive input $p(x_{true}|x_{sensed})$) by taking into account the likelihood distributions of the prior (i.e., the cumulative information from prior experience) and the evidence (i.e., the current proprioceptive input x_{sensed}):

$$p(x_{true}|x_{sensed}) = p(x_{sensed}|x_{true}) * \frac{p(x_{true})}{p(x_{sensed})} \quad (7)$$

where $p(x_{sensed}|x_{true})$ is the likelihood of a particular proprioceptive input x_{sensed} when the perceived body state really is true. This then allows the inference of the current body state given past experience and incomplete or polluted proprioceptive inputs.

Law of Continuity

Objects and points that are co-linear and follow the same direction will be grouped together as a whole (see Figure 2d). We can construct proficient continuations between neighboring local environments. Density-Based Spatial Clustering of Applications with Noise (DBSCAN) identifies outliers as noises. The Mean-shift algorithm, Eq. (8), actually includes them in the cluster despite differences of the data point. DBSCAN also does not require a pre-set number of clusters, and discovers arbitrarily shaped clusters. These are key facets for analysing proprioceptive data.

Given the manifolds of afferent information for natural movements are usually continuous, then the Law of Continuity would naturally apply as the manifold during a movement continues along a particular path, even if temporarily disrupted or occluded by a perturbation. In practice, Bayes' Rule is a way in which such expectation of continuity can be quantified.

Law of Prägnanz (Good Form, Clarity)

The Law of Prägnanz focuses on simplicity and will prompt visualizations according to the simplest way of grouping items. We perceptually organize shapes to simple forms, as in pithiness. The Law of Prägnanz is the tendency to interpret ambiguous images as simple and complete vs. complex and incomplete. An example is how shapes overlapping each other can cause ambiguity, as shown in Figure 2e. A potential resolution is an iterative method such as Mean-shift Clustering, Eq. (8), where $N(x)$ is the neighborhood of the set of points, x . Depending on the Gaussian kernel bandwidth, Eq. (9), the Mean-shift algorithm iteratively shifts points until there is a convergence of partitioning the clusters into semantically meaningful groups. This is probable to work well with proprioceptive afferents as it may account for the noise in signals which is expected, and necessary for physiological function.

$$m(x) = \frac{\sum_{x_i \in N(x)} K(x_i - x)x_i}{\sum_{x_i \in N(x)} K(x_i - x)} \quad (8)$$

$$K(x_i - x) = e^{-c\|x_i - x\|^2} \quad (9)$$

Dimensionality reduction is probably the most commonly applied (and potentially misinterpreted) analysis of

high-dimensional motor signals (Kutch and Valero-Cuevas, 2012). It is simply a way to quantify whether a high-dimensional ensemble of signals evolves (i.e., has variance) along all dimensions equally, or inhabits a lower-dimensional subspace. Conceptually, it is just the singular value decomposition of a covariance matrix, where the number of 'large' singular values (principal components) quantifies the rank of the covariance (the effective 'dimensionality' of the data), and the left singular vectors (principal vectors) form a basis for those dominant variances (the basis for the effective subspace the data inhabit). Independent Component Analysis and Nonnegative Matrix Factorization is a variations on this idea that do not require orthogonality of the basis vectors, and the latter also imposes a non-negative constraint on the elements of the basis (as neural signals are conceptualized as intensities or spiking frequencies that are > 0). It is good to see that some work is beginning to be done on dimensionality reduction in tactile afferents, which are famously difficult to record from even in animal preparations (Rongala et al., 2018). Our current work is beginning to apply dimensionality reduction to higher-dimensional simulated proprioceptive signals Berry et al. (2020).

Supplementary Laws

Other Gestalt grouping laws that can be applied to sensory stimuli integration include the Laws of Focal Point, Symmetry, Common Fate, Common Region, Synchrony, Convexity, Isomorphism, Parallelism, Unity, Element Connectedness, and Figure vs. Ground.

Functional Utility of the Sensory-Motor Gestalt

Figure 4 describes our working hypothesis of the Sensory-Motor Gestalt in operation. At any time point, proprioceptive (and other sensory) information define a state of the body that lies within a particular manifold of like inputs (Laws of Proximity and Similarity) and feasible next states (Laws of Continuity and Closure). Such body state allows feasible transitions to 'next' proprioceptive states via feasible motor actions that will lead to a, usually predicted, new body state (Laws of Continuity and Closure).

Abstracting Self from Sensorimotor Experiences for Neuromuscular Systems

Now let's examine how the foundations of neuromuscular systems can provide context to constructing the minimal operative self via proprioceptive signals. In Nature, proprioception provides animals with awareness of the state of their body and of their relation to the environment. Proprioceptive signals arise from mechanoreceptors that reflect the state of tissues, which are driven by muscle forces, joint and body postures, and skin deformations. When integrated with other

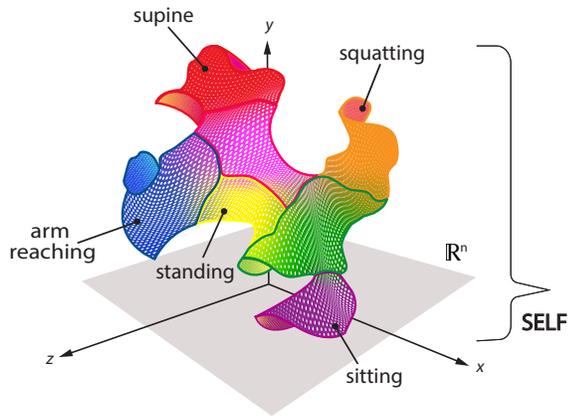


Figure 4: We envision the representation of minimal self as a collection of categorized states in \mathbb{R}^N space formed from sensory and motor maps, and made useful by the agency they provide. Our data-driven projection method categorizes the set of feasible inputs from muscle spindles for each specific task performed (i.e., arm reach, sit, squat, standing) as a manifold. Transitions from one state to another occur through point-to-point transitions along the manifold. The high-dimensional space of afferent modalities has an underlying structure given by the anatomy of the body and the physical transitions it can undergo such as changing postures via self movement.

sensory modalities, this reflection of body state at any given moment in time and space provides the nervous system with an overall representation of bodily position, actions, and task experiences. Neuroscientists have long been intrigued with how the brain represents the body and forms models of bodily states through proprioception (Graziano, 2000). However, there is still no consensus regarding how these representations, facilitated by multi-muscle control, compartmentalize and process high-dimensional afferent information as

continuous feedback for ongoing tasks.

The fundamental formulation of a control law for a linear system (without loss of generality) is

$$\dot{x} = Ax + Bu \quad (10)$$

where the outputs y (and therefore sensory and proprioceptive signals) are a function of the state x and the control signals u

$$y = Cx + Du \quad (11)$$

By definition, the equations of motion (i.e., $\dot{x} = Ax$) are an important determinant of the feasible transitions away from any given state. Moreover, changes in sensory and proprioceptive signals are driven by changes in state (i.e., $y = Cx$). This is a formal way to conceptually anchor some aspects of the Sensory-Motor Gestalt. *Please note we do not claim or endorse that the engineering concept of ‘state’ applies to biology. But the Sensory-Motor Gestalt is a formal way to describe how the stream of sensory and proprioceptive signals is useful to biological behavior in a way that is agnostic to how those signals are processed.*

We can conjecture how the nervous system processes incoming afferents (e.g., proprioceptive signals) by observing how neural activation commands mathematically map to mechanical outputs, as shown in Figure 5. Neural commands simply refer to the nervous system’s distribution of excitatory impulse signals to activate muscle tissue. For tendon-driven limbs, Valero-Cuevas (2016) emphasizes that the nervous system’s primary function is to use (i.e., learn, explore, and exploit) the set of feasible neural commands from the optimized activation space with dimensionality of vector $\mathbf{a} \in \mathbb{R}^N$, where N is the number of independently controlled muscles. From activation space, vector spaces are successively mapped to muscle force space, to joint torque space, then lastly to output wrench space to produce a set of feasible mechanical outputs (forces and movements).

In prior work (Berry et al., 2017), we have extrapolated this perspective of muscle redundancy to feasible sets of proprioceptive signals, called Feasible Sensory Sets (FSS).

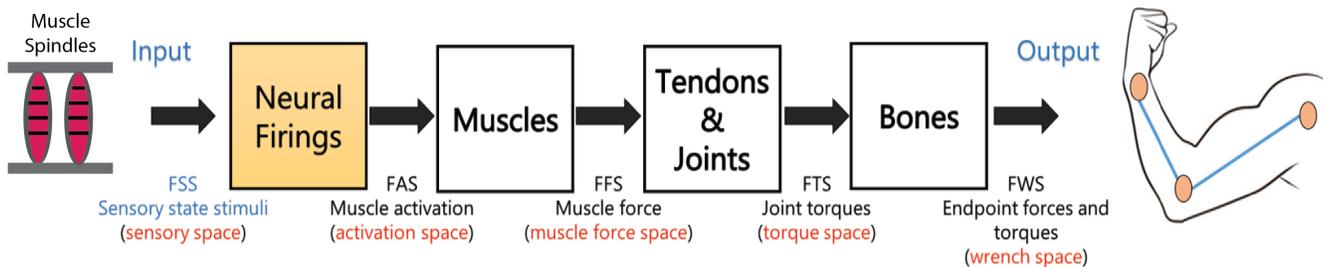


Figure 5: The neuromechanical perspective of how sensory inputs are transformed to motor outputs (adapted from Valero-Cuevas (2016)). A Feasible Sensory Set (FSS) defines the afferent stimuli that are plausibly detectable for a given state of the body (i.e., joint posture, force production, and kinematic task). By incorporating the influence of proprioceptive space via neural spike firings, an under-constrained mapping of transformations can be reinforced from neural motor commands in the Feasible Activation Sets (FAS) to mechanical outputs (limb movements).

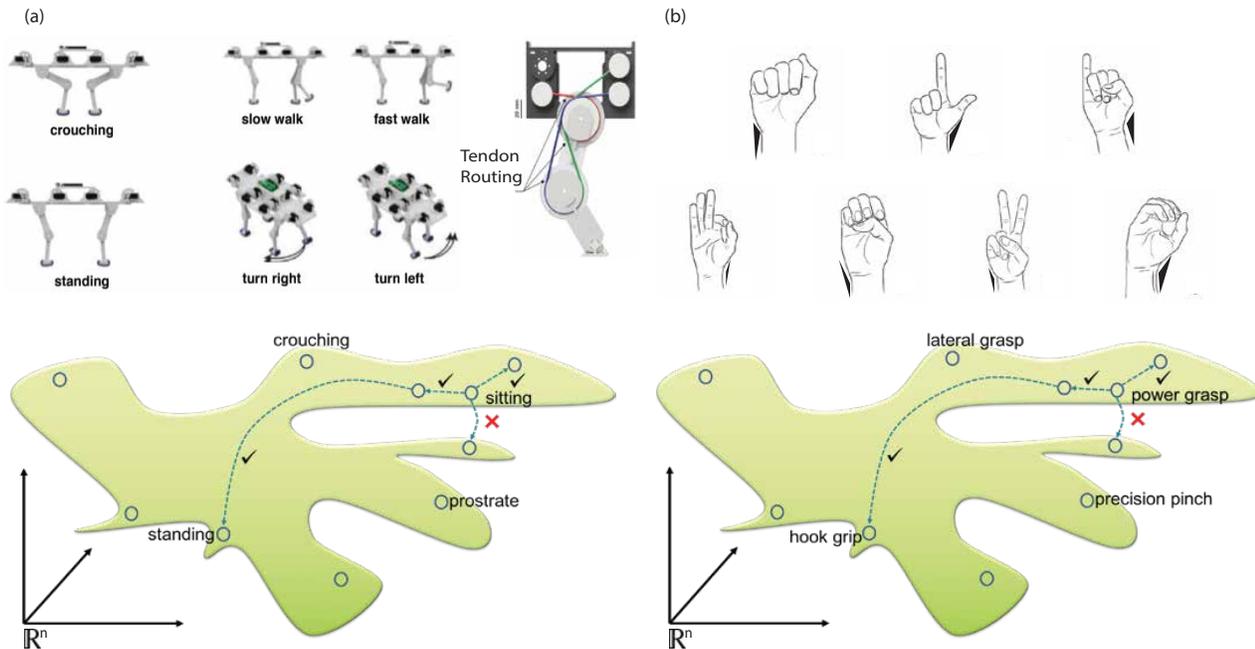


Figure 6: Test-bed applications for Sensory-Motor Gestalt implementation in tendon-driven systems. The Gestalt can provide state cases and their feasible transitions along the manifold for agents. (a) Proposed neuromorphic cat-like robot and limb in hardware. Image adapted from Marjaninejad et al. (2019). (b) Various hand and finger states depicted using the American Sign Language. The nervous system registers the proprioceptive feedback generated with each hand shape is unique to each letter signed with gestures.

These are defined by a body’s anatomical structure and the mechanical tasks being performed. Here, we first introduced the concept of trajectory-specific proprioceptive manifolds, which are the unique multidimensional and time-varying combinations of afferent signals that obligatorily emerge during a limb movement. We demonstrated that a given movement gives rise to a distinct sensory manifold embedded in the 12-D space of spindle information that is largely independent of the choice of muscle coordination strategy. These are referred to as manifolds because they are a systematic collection of points (i.e., spindle neural spikes) that provide information for its control.

Following this work it remained unknown whether spindle signals suffice to discriminate limb movements. We used a 4-musculotendon, 2-joint cat hindlimb model to simulate muscle spindle length and velocity signals (II and Ia, respectively) during repeated cycles of five distinct endpoint movements, similar to the manifolds in Figure 3. In Berry et al. (2020), we concluded that proprioceptive information can usefully discriminate limb states—but only after conducting minimal pre-processing of high-dimensional multi-muscle ensembles to low-dimensional subspace components. This finding may this explain the documented subcortical pre-processing of afferent signals of various mammals (Rongala et al., 2018). It is this resulting set of constrained sensory signals that we believe could suffice as a minimal representation of the artificial self and should be

incorporated into the Sensory-Motor Gestalt paradigm. We project the usefulness of Sensory-Motor Gestalt to be a suitable core to execute on different applications that utilize neuromuscular dynamics, incorporate neuromorphic and bio-inspired architectures, and classification of human bodily states (Figure 6).

Role of Self and Identity in Autonomous Robotics and Synthetic Biological Agents

A semblance of selfhood, identity, and agency should be expected outcomes for constructing a dynamic sensorimotor representation (Thompson, 2005). For robots, concepts of identity are typically viewed as a necessity for interactions in social environments (Duffy, 2004). For humans, the self and identity combination are purposed for storing the traits, stereotypes, characteristics, and roles they play in social settings. What features constitute a person’s self? How do disparate sensory perceptions cohesively fuse together to form a singular experience of self? Although these questions are typically addressed within the human scope, we can also apply these inquires to autonomous robots that are bioinspired and create their own experiences with action.

It is our opinion that sensorimotor contingencies (discussed in Related Work) do not achieve their full potential if solely used to estimate error signals in closed-loop controllers. We believe one can ask the extent to which these contingencies facilitate and embody a self, reflect an iden-

tity, and activate agency needs to be thoroughly explored. Self and identity are often used interchangeably to encapsulate the entirety of individual's behavior, character, and the restricted contextual constraints in which they operate within. However, it is important to clearly know the distinctions of these terms if we're determined to adequately construct models that emulate their functions.

Agency is known as the control of intentional actions and volition; leading to the ability to plan and action ownership. For the purpose of our study we distinguish the self and identity according to Oyserman's (2012) conceptualization. It is thought best to consider self, self-concept, and identity as nested elements: self is the top-tier construct, self-concepts reside within the self, and identities reside within self-concepts. Oyserman defines self as the ability to consider oneself as an object. The self maintains reflexive capacity that is able to direct an agent to what is "me"; it is the focal point of personal account and a reference for anchoring temporal sequences of events (e.g., memory recall). Identities are "content and readiness to act and employ mindsets to make meaning." Personal identities are the traits, characteristics, values and goals belonging to the agent. Altogether self and identity are mental concepts, social products, and forces of action. As Oyserman states, what makes this nested unit interesting is that they appear to predict behavior over time. What is not fully understood by many in literature is *how* this happens.

Related Work

Further research into the topic of dynamic sensorimotor representations led us to original work on the sensorimotor contingency theory (O'Regan and Noë, 2001), which has motivated an assortment of studies in the area of human perception as it relates to understanding the nature of actions and their sensory effects. Sensorimotor contingencies derived from the notion that vision should be treated as an environmental exploratory activity. According to Hay et al. (2018), sensorimotor contingencies spawned multi-disciplinary projects that investigated how to model the action-sensory relationship of robotic systems, which spanned the manipulation, classification, and categorization of external objects. The researchers view the goal for most of these studies as autonomous robots learning skilled behaviors via learning the structure of complex sensorimotor spaces and how actions affect the environment. Despite these contributions, Buhrmann et al. (2013) believed there have been few attempts to formally define sensorimotor contingencies, which they view as a prerequisite for testing this approach via models and empirical study. The sensorimotor contingencies view on perceptual awareness have also been criticized for lacking a suitable foundation in the biology of autonomous agency. Prior work on building computational approaches of body representations, self, identity and Gestalt have been attempted with a grounding in min-

imal embodied, psychological and cognitive aspects (Gallagher, 2013). One drawback of past implementations is that they're unencumbered with understanding the manifold of feasible transitions, which therefore leads to the incorrect perspective that any action is permissible. Our approach addresses how the sensorimotor self should constrain one's agency and perceptual space to feasible tasks. Attempts to create sensory-to-motor maps as a body representation (i.e., body schema) have been accomplished by achieving robotic self-recognition using a dynamic Bayesian network (Gold and Scassellati, 2009), online learning of arm reaching motor maps for humanoid robots using open and closed loop control (Gaskett and Cheng, 2003), body representations as cross-modal map learning of invariance in multi modal sensory data (Yoshikawa et al., 2003), and estimation of a kinematic model for serial robots (Martinez-Cantin et al., 2010). To our knowledge, our approach is the first that considers the inherent link among the feasible capabilities of the body, the feasible sensory information that will emerge, and the physics of the world as the manifold defining agency, and therefore delineating the concept of self.

Conclusion

The emergence of self and its role in biological and artificial agents continues to be a subject of debate across many disciplines; leading to the perception that there is a lack of congruence among perspectives. The absence of a unified concept of self presented us with an opportunity to propose a sensorimotor mechanism by which the self can emerge, via the Gestalt laws of perceptual organization, in the context of artificial systems operating in the physical world (i.e., robots). This enables us to investigate the foundation of self, identity, learning, and agency as the multifaceted interplay of proprioception and action while exploring their implications to autonomy. The emergence of self through sensorimotor interactions has applications ranging from a self-other distinction to 'social' systems for robot-human and robot-robot interactions. Traditionally, self and identity are considered to be theoretical concepts, social constructs, and therefore enablers of agency. We visit these concepts in reverse order to propose that it is sensorimotor agency that can enable the emergence of self and identity—which is an evolutionarily plausible order of events (Wilson, 1999). The Sensory-Motor Gestalt provides a solid foundation to enable such cross-fertilization to move towards the creation of truly autonomous and versatile robots, and promote advances in artificial intelligence.

Acknowledgements

This work is supported in part by the NIH (R01 AR-050520 and R01 AR-052345), DOD CDMRP Grant MR150091, DARPA-L2M program Award W911NF1820264, and the USC Viterbi School of Engineering. This work does not necessarily represent the views of the NIH, DoD, or DARPA.

References

- Berry, J., Marjaninejad, A., and Valero-Cuevas, F. J. (2020). Minimal pre-processing of multi-muscle ensembles of spindle signals improves discriminability of limb movements. In Preparation.
- Berry, J. A., Ritter, R., Nagamori, A., and Valero-Cuevas, F. J. (2017). The neural control of movement must contend with trajectory-specific and nonlinearly distorted manifolds of afferent muscle spindle activity. In *IJCNN*, pages 1188–1194. IEEE.
- Boyer, D. M., Lipman, Y., Clair, E. S., Puente, J., Patel, B. A., Funkhouser, T., Jernvall, J., and Daubechies, I. (2011). Algorithms to automatically quantify the geometric similarity of anatomical surfaces. *Proceedings of the National Academy of Sciences*, 108(45):18221–18226.
- Buhrmann, T., Di Paolo, E. A., and Barandiaran, X. (2013). A dynamical systems account of sensorimotor contingencies. *Frontiers in psychology*, 4:285.
- Chang, D. H.-J. (2007). The gestalt principles of similarity and proximity apply to both the haptic and visual grouping. In *Eighth Australasian User Interface Conference (AUIC2007)*, pages 79–86. Flinders Press.
- Duffy, B. (2004). Robots social embodiment in autonomous mobile robotics. *International Journal of Advanced Robotic Systems*, 1(3):17.
- Faivre, N., Arzi, A., Lunghi, C., and Salomon, R. (2017). Consciousness is more than meets the eye: a call for a multisensory study of subjective experience. *Neuroscience of consciousness*, 2017(1):nix003.
- Gallagher, S. (2013). A pattern theory of self. *Frontiers in human neuroscience*, 7:443.
- Gaskett, C. and Cheng, G. (2003). Online learning of a motor map for humanoid robot reaching. *Proc. 2nd Int. Conf. Computat. Intell., Robot. Autonom. Syst. (CIRAS 2003)*.
- Gold, K. and Scassellati, B. (2009). Using probabilistic reasoning over time to self-recognize. *Robotics and autonomous systems*, 57(4):384–392.
- Graziano, M. S. (2000). How the brain represents the body: insights from neurophysiology and psychology. *Common mechanisms in perception and action*.
- Guarino, N., Oberle, D., and Staab, S. (2009). What is an ontology? In *Handbook on ontologies*, pages 1–17. Springer.
- Hay, N., Stark, M., Schlegel, A., Wendelken, C., Park, D., Purdy, E., Silver, T., Phoenix, D. S., and George, D. (2018). Behavior is everything: Towards representing concepts with sensorimotor contingencies. In *Thirty-Second AAAI Conf.*
- Körding, K. P. and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244–247.
- Kubovy, M., Holcombe, A. O., and Wagemans, J. (1998). On the lawfulness of grouping by proximity. *Cognitive psychology*, 35(1):71–98.
- Kutch, J. J. and Valero-Cuevas, F. J. (2012). Challenges and new approaches to proving the existence of muscle synergies of neural origin. *PLoS computational biology*, 8(5).
- Marjaninejad, A., Urbina-Meléndez, D., Cohn, B. A., and Valero-Cuevas, F. J. (2019). Autonomous functional movements in a tendon-driven limb via limited experience. *Nature machine intelligence*, 1(3):144–154.
- Martinez-Cantin, R., Lopes, M., and Montesano, L. (2010). Body schema acquisition through active learning. In *2010 IEEE international conference on robotics and automation*, pages 1860–1866. IEEE.
- O’Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, 24(5):939–973.
- Oyserman, D., Elmore, K., and Smith, G. (2012). *Self, self-concept, and identity. Handbook of self and identity.*, pages 69–104.
- Rongala, U. B., Spanne, A., Mazzoni, A., Bengtsson, F., Oddo, C., and Jörntell, H. (2018). Intracellular dynamics in cuneate nucleus neurons support self-stabilizing learning of generalizable tactile representations. *Frontiers in cellular neuroscience*, 12:210.
- Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the cognitive sciences*, 4(4):407–427.
- Valero-Cuevas, F. J. (2016). *Fundamentals of neuromechanics*. Springer.
- Warshall, S. (1962). A theorem on boolean matrices. *Journal of the ACM (JACM)*, 9(1):11–12.
- Wertheimer, M. (1923). Laws of organization in perceptual forms. *A source book of Gestalt Psychology*.
- Wilson, F. R. (1999). *The hand: How its use shapes the brain, language, and human culture*. Vintage.
- Yoshikawa, Y., Hosoda, K., and Asada, M. (2003). Does the invariance in multi-modalities represent the body scheme?-a case study with vision and proprioception. In *2nd Intelligent Symposium on Adaptive Motion of Animals and Machines*. Cite-seer.