

# Perceptual Spaces: Mathematical Structures to Neural Mechanisms

Qasim Zaidi,<sup>1</sup> Jonathan Victor,<sup>2</sup> Josh McDermott,<sup>3</sup> Maria Geffen,<sup>4</sup> Sliman Bensmaia,<sup>5</sup> and Thomas A. Cleland<sup>6</sup>

<sup>1</sup>Graduate Center for Vision Research, State University of New York College of Optometry, New York, New York 10036, <sup>2</sup>Brain and Mind Research Institute, Weill Cornell Medical College, New York, New York 10021, <sup>3</sup>Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, <sup>4</sup>Department of Otorhinolaryngology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, <sup>5</sup>Department of Organismal Biology and Anatomy, University of Chicago, Chicago, Illinois 60637, and <sup>6</sup>Department of Psychology, Cornell University, Ithaca, New York 14853

A central goal of neuroscience is to understand how populations of neurons build and manipulate representations of percepts that provide useful information about the environment. This symposium explores the fundamental properties of these representations and the perceptual spaces in which they are organized. Spanning the domains of color, visual texture, environmental sound, music, tactile quality, and odor, we show how the geometric structures of perceptual spaces can be determined experimentally and how these structures provide insights into the principles of neural coding and the neural mechanisms that generate the codes, and into the neural processing of complex sensory stimuli. The diversity of the neural architecture in these different sensory systems provides an opportunity to compare their different solutions to common problems: the need for dimensionality reduction, strategies for topographic or nontopographic mapping, the utility of the higher-order statistical structure inherent in natural sensory stimuli, and the constraints of neural hardware.

## Introduction

The energy absorbed by sensory receptors provides organisms with clues about the structure of the world, but the properties of the world that are important to an organism are rarely explicit in the sensory input. One of the main aspirations of neuroscience is to understand how populations of neurons build and manipulate representations that make explicit useful properties of the world. This symposium aims at generating interest into the fundamental properties of these representations, known as perceptual spaces, across multiple sensory modalities.

A perceptual space consists of a set of stimuli in some sensory domain along with a set of similarity relationships. The characteristics of a perceptual space center on two fundamental properties: dimensionality and intrinsic geometry. Dimensionality and intrinsic geometry are, in turn, consequences of the space's metric (i.e., the operation that defines similarity). Behaviorally, similarities can be estimated by matches (including generalization), just-discriminable differences, proximity to exemplars, and mid-point settings. Physiologically, similarities can be quantified by activation patterns of receptors or other neurons, in rates

or temporal patterns of impulse responses, and in different levels of correlated firing.

Based on the experimentally determined properties of the similarity measure, the perceptual space can be given a well-defined geometry, thus providing access to a large number of theorems that in turn specify implications of the representational structure. These geometries form a natural hierarchy, with more highly structured geometries placing greater demands on the conditions that the metric must satisfy (Klein, 1939; Brannan et al., 1999). At the top of the hierarchy is familiar Euclidean geometry and its non-Euclidean relatives, which allow representing stimuli as vectors, with well-defined sizes and angles. Affine geometry is one step down the hierarchy: it allows for vector representations on arbitrarily scaled axes and has a notion of lines and parallelism, but no notion of angle or size. A further step down is projective geometry: collinearity and dimension remains defined, but not parallelism. At the bottom of the hierarchy (with the fewest geometric requirements) is the notion of a topological space: proximity is defined, but collinearity and dimension need not be. Superimposed on this characterization of the intrinsic geometry of the perceptual space is its extrinsic geometry, which is the mapping of the perceptual space on a physical space of stimuli; characterizing this can provide additional information about neural transformations. Higher-dimensional representations enable added flexibility in learning and finer-grained qualitative distinctions but can impose a higher cost on similarity computations.

Maxwell's studies (Maxwell, 1860) provide a paradigmatic example of the analysis of a perceptual space. By showing that color matches satisfied the linearity properties of additivity and scalar

Received Aug. 5, 2013; revised Sept. 13, 2013; accepted Sept. 18, 2013.

Author contributions: Q.Z., J.V., J.M., M.G., S.B., and T.A.C. wrote the paper.

This work was supported by National Institutes of Health Grants EY07556 and EY13312 to Q.Z., National Institutes of Health Grants EY 07977 and 09314 to J.V., McDonnell Scholar Award to J.M., Burroughs Wellcome Career Award, Klingenstein Fellowship, and PA Lions Hearing Research Grant to M.G., National Science Foundation Grant IOS-1150209 to S.B., and National Institutes of Health Grant DC012249 to T.A.C.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Qasim Zaidi, SUNY College of Optometry, Graduate Center for Vision Research, 33 West 42nd Street, New York, NY 10036. E-mail: qz@sunyopt.edu.

DOI:10.1523/JNEUROSCI.3343-13.2013

Copyright © 2013 the authors 0270-6474/13/3317597-06\$15.00/0

multiplication, Maxwell was able to embed color matches into the structure of the linear algebra recently invented by Grassman, thus allowing for vector operations to predict the results of combining lights of different colors. Moreover, although the physical combinations of visible lights that range from 400 to 700 nm form a space of arbitrarily high dimensionality (depending on the discrete grain of dividing the spectrum), Maxwell showed that the space of color matches was only 3D: the key observation was that for color-normal observers, any light could be matched by a suitable combination of any three fixed primary lights. To probe the physiological basis of perceptual color space, Maxwell went on to show that the matches of color-normal observers were a reduced subset of the matches of congenital color defectives. He then used the additional color confusions of individuals with congenital color vision deficits to derive the spectral sensitivities of the three types of cones. A century later, these predictions were confirmed: electrophysiological measurements of the spectral sensitivities of human cones (Schnapf et al., 1987) were shown to match psychophysical estimates derived using Maxwell's methods (Smith and Pokorny, 1975). Variants of Maxwell's space, for example, CIE 1931 or MacLeod and Boynton (1979) have proven invaluable for perceptual and physiological investigations of the visual system, and Maxwell's notion of metamers (physically distinct stimuli that appear identical) has proven to be an important concept in other sensory domains (Freeman and Simoncelli, 2011).

Modern computational and experimental developments in manipulating complex stimuli enable the construction of perceptual spaces that can provide insights into the shaping of peripheral and central neural tuning by the characteristics of natural stimuli, the roles of temporal and spatial factors in neural coding, the neural basis of salience and perceptual scale invariance, and learning processes that embody abstraction, categorization, and generalization. To exemplify these insights, the presentations in this symposium discuss the neural representations of colors, natural images, environmental sounds, tactile materials, and natural odors. Despite the diversity of the neural architecture in these domains, the problems confronted by sensory systems have much in common: the need for dimensionality reduction, the choice between topographic and nontopographic mapping, the opportunity to use the statistical structure inherent in natural sensory stimuli, and the constraints of neural hardware. These examples provide an opportunity to compare and contrast perceptual spaces and their underlying neural processes across these modalities.

### Geometric structure of perceptual color space

Historically, color vision has seen the most systematic development of perceptual spaces for research and applied purposes, so it is an appropriate starting point for this symposium. Although most investigations of color processing focus on color matching and discrimination, the use of color to make inferences about the environment, and guide action, requires more sophisticated processing, such as the use of relative color similarities to identify materials across spectrally distinct illuminations (Zaidi, 1998; Zaidi and Bostic, 2008). Multidimensional scaling (Shepard, 1962) has been used to specify color spaces based on numerical ratings of similarity between colors, but the procedure requires Euclidean assumptions, and Wuerger et al. (1995) used proximity judgments between points and lines to show that this assumption fails for color similarity. Thus, the intrinsic geometry of color space is as yet unclear: it must be structured enough to support judgments of relative similarity, but proximity judgments indicate that it is not Euclidean.

Given these results, Ennis and Zaidi (2013) directly investigated the intrinsic geometric structure supporting a color similarity space. Varignon's Theorem (Coxeter and Greitzer, 1967) states that, for any quadrilateral, when the midpoints of opposite sides are joined by straight lines, the point of intersection is simultaneously the midpoint of both lines. This theorem holds only for vector spaces that are at least affine, so they tested its validity for perceived midpoints of quadrilaterals covering vastly different colors. Observers viewed a test patch flanked by two patches, each containing one vertex color of a quadrilateral's edge. They were instructed to consider the color change between the endpoints along the "reddish-greenish" and "bluish-yellowish" opponent-dimensions and to adjust the middle patch's hue and saturation to a color perceived as the midpoint on both dimensions (i.e., equally similar to both endpoints). After finding the midpoints for the four sides, observers set the midpoints between the two pairs of facing midpoints. For seven observers, the two final midpoints for each quadrilateral coincided, thus satisfying the conditions for an affine space. Therefore, when observers explicitly use an opponent-color mental representation, a perceptual color space based on relative similarities across large color differences has an affine structure. In a Euclidean color space, the distance between chromaticities would represent magnitude of similarity, but even in a weaker affine space, ratios of distances along a line would provide measures of relative similarity, and parallelism would provide similarity between color changes.

The usefulness of similarity judgments is not confined to judging materials across illuminations; they are also used in every sensory modality for a number of tasks. Indeed, similarity is one of the fundamental principles used by Gestalt psychology to explain perceptual organization phenomena (Wertheimer, 1912) and is essential to generate models of generalization (Tversky, 1977; Shepard, 1987). The neural basis of similarity computations is essentially unknown, partly because of the paucity of behavioral tasks that could be used with a model animal and the lack of empirical justification of the geometrical properties. The color results show a path toward ameliorating the situation. A midpoint between two stimuli is the stimulus that is simultaneously most similar to the two and could be ascertained by having an animal do forced choices of similarity to first find the locus of stimuli that are equally similar to two fixed stimuli and then from this set the stimulus that is most similar to both. If the perceptual space of the animal could be shown to be Euclidean or affine, it puts strong constraints on the neural circuits for computing similarity: Euclidean distances imply a comparison based on the power (sum of squares) of the difference; affine geometry implies comparisons based on ratios.

### Perceptual space of the elements of spatial vision

By investigating the perceptual space of elements of spatial vision, Victor and Conte (2012) explored how the characteristics of early visual processing are shaped by a combination of the characteristics of natural sensory stimuli and the constraints of neural hardware.

Random images containing the elements of spatial vision (light vs dark, lines, edges, corners, etc.) can be considered to constitute a perceptual space, but the coordinates and geometry of this space are unclear, as these elements are interdependent. For example, edges typically occur when luminance differences are present, and corners typically occur only when lines or edges are present. To navigate in this space, Victor's laboratory first constructed a large set of artificial images that allowed for the

independent manipulation of these elements, up to the constraints of their mathematical dependencies. The resulting stimulus space was 10-dimensional: rich enough to capture the geometric interdependence of local image features but still tractable (Victor and Conte, 2012).

Working within this space, Victor and Conte (2012) determined its perceptual structure (i.e., the extent to which images defined by different complements of local features were visually distinct). They found that the perceptual structure could be described by an “isodiscrimination surface,” the set of points in the space that were equally distinguishable from the random image. The isodiscrimination surface had an ellipsoidal shape, indicating that the local geometry of the perceptual space was Euclidean. The axes of the ellipsoid, which represented the features that were most visually salient, next-most-salient, etc., did not correspond to the “pure” features, such as luminance or edges at single orientation; rather, each perceptual axis was a combination of several of these elements. The analysis also showed that some combinations of features were highly salient whereas others were virtually imperceptible.

In collaboration, Briguglio et al. (2013) determined the distribution of these same local elements in natural images. Interestingly, this showed that the pattern of covariation of these local elements in natural images was a close match to the pattern of visual sensitivity. In other words, the axes in the perceptual space of greatest sensitivity corresponded to the combinations of local image statistics that were maximally informative, indicating a kind of detailed matching of early visual processing with the characteristics of natural visual stimuli.

### Insight into the perceptual space of sound from the analysis and synthesis of real-world audio

The auditory system receives as input the pattern of displacement of the two eardrums, which vibrate back and forth in response to sound. From that pattern of displacement, an organism must determine what happened in the world: whether it is listening to an attractive conspecific, an upset baby, or a rainstorm. In most cases of interest, it is computationally intractable to determine the nature of a sound source directly from the sound waveform. Instead, sensory input must be transformed through layers of neural circuitry into representations in which the properties of interest are more readily discernible. Such representations form the basis of perceptual spaces; they highlight behaviorally important dimensions of variation, such as those that distinguish different classes of stimuli, and collapse across sources of variation that are irrelevant for behavior.

To test theories of such perceptual representations, one approach is to test whether a candidate representation can support discrimination between classes of stimuli (Pinto et al., 2009). An alternative approach is to evaluate a model representation by synthesizing stimuli that generate particular values in the representation. If the representation discards (and retains) the same information that the perceptual system does, synthetic examples that produce the same representation as some real-world stimulus, such as a musical instrument sound, ought to be perceived as another realistic example of the same thing (Risset and Wessel, 1999; Gaver, 1993; Portilla and Simoncelli, 2000).

McDermott and colleagues (McDermott et al., 2009; McDermott and Simoncelli, 2011) recently adopted this approach to the study of sound textures: sounds produced by the superposition of many similar acoustic events, as a result from falling rain, swarms of insects, or galloping horses. Different textures vary in their statistical properties, suggesting that textures could be repre-

ented with statistics computed from the peripheral auditory system. McDermott and colleagues (McDermott et al., 2009; McDermott and Simoncelli, 2011) examined statistics, such as the mean and variance of the responses of simulated auditory filters, as well as correlations of various sorts between different filter responses. Because such statistics are averages of information over time, they are potentially suited to characterizing stationary properties of a sound signal (e.g., the intensity of a rainstorm, or the number of people clapping in an audience). Consistent with this notion, random signals constrained only to have the same values of particular statistics as a real-world texture often sounded like realistic examples of the texture (McDermott et al., 2009; McDermott and Simoncelli, 2011). This result gives credence to the hypothesis that textures are represented with relatively simple statistics computed from the responses of the peripheral auditory system.

Notably, different examples of the same texture (each with a distinct arrangement of details) have the same statistics, such that a representation consisting of the statistics provides a form of invariance. This observation raises an interesting possibility: if the perceptual space of texture is defined solely by these statistics, listeners might have difficulty distinguishing different examples of the same texture. This is indeed the case for moderately long excerpts (more than a few seconds in length): different examples of the same texture are difficult to discriminate, even though excerpts of different textures (with different statistics) are readily distinguished (McDermott et al., 2013). Textures thus appear to be represented in a space of sound statistics. The acoustic events composing the texture (e.g., crackles, claps) are incorporated into statistical measurements but are not themselves readily accessible to the listener.

### Population code for invariant representation of natural sounds in the mammalian auditory system

Discriminating and detecting complex sounds are complicated in the real-world environment by acoustic distortions of the communication signals when produced by different sources and in the presence of background noise. The sound representation at the early auditory processing stages reflects the acoustics of the signal and is therefore expected to change dramatically with acoustic distortions or background noise (Pfafflin and Mathews, 1966; Bregman, 1990; Isabelle and Colburn, 1991). To comprehend a word spoken by different speakers or in different noise environments, the auditory system must develop invariant representation of complex sound objects, by converting the representation of the physical features of the sound waveform to an object-based representation. According to the hierarchical processing model, in higher brain areas, representation of complex sounds, such as vocalizations, is expected to change less with the vocalizations' acoustic distortions or background noise than in earlier auditory areas, as neuronal population exhibits increased invariance (Bregman, 1990; Russ et al., 2007; Nelken and Bar-Yosef, 2008).

Mixed evidence exists on whether the primary auditory cortex (A1), a key auditory processing area in mammals, exhibits invariance to basic temporal and spectral transformations of con-specific vocalizations. Previously, Geffen and colleagues found that individual neurons in A1 exhibit preferential representation of the original, over distorted, vocalizations. Next, Geffen and colleagues tested whether populations of neurons in A1, recorded in awake rats, exhibit invariance to transformations in the basic acoustic features of con-specific vocalizations. To (Carruthers et al., 2013) evaluate how the population code of primary auditory

cortex facilitates invariant representations of similar stimuli, simultaneous neuronal activity from ensembles of neurons in A1 was recorded in response to a set of vocalizations, presented both as original recordings and after amplitude- and frequency-modulation rates had been increased or decreased. Classification accuracy of support vector machines, which were fit and tested on each stimulus condition, was significantly higher when tested on the same stimulus condition on which it was trained. However, the performance of the classifier, when trained to and tested on different conditions, was significantly above chance. Thus, although not demonstrating invariant representation directly, the population code in A1 contains enough information for a circuit just one synapse up to construct an invariant representation of temporally and spectrally distorted signals. Furthermore, population-level analysis of the mechanisms used to extract the representation of transformed and original vocalizations demonstrates that noise correlations between neuronal responses play a crucial role. Noise correlations between neurons improved the accuracy of classification of vocalizations. Downstream neuronal circuits likely use information encoded in synchronous activity between A1 neurons.

Combined, these results demonstrate that invariant representation of complex sounds in the mammalian auditory system begins at the level of populations of neurons in the primary auditory cortex. Therefore, the representation of the acoustic perceptual space in the brain can only be understood at the level of population codes.

### Spatial and temporal mechanisms for tangible textures across spatial scales

The somatosensory system supports a rich perceptual space of surface textures, comprising at least four major dimensions: roughness/smoothness, hardness/softness, stickiness/slipperiness, and warm/cool (Hollins et al., 2000). Although surface texture itself is intrinsically spatial, the neural mechanisms underlying the perception of surface texture rely on a combination of spatial and temporal codes. Coarse surface features (on the order of millimeters) are reflected in the spatial pattern of activation evoked in a population of mechanoreceptive afferents that innervate the fingertip skin (Connor and Johnson, 1992) in a relatively dense manner ( $\sim 100$  units/cm<sup>2</sup>). However, much of the “feel” of a surface is driven by features on the order of  $1\ \mu\text{m}$ , a much finer scale. To sense these features, we run our fingers across the surface, which leads to the elicitation of small and complex skin vibrations (Hollins et al., 2002; Bensmaia and Hollins, 2003). Two populations of cutaneous mechanoreceptive afferents are exquisitely sensitive to such skin vibrations and produce highly temporally patterned and repeatable responses to textures (Mackevicius et al., 2012; Weber et al., 2013). The spike patterns are texture-specific and contain enough information to mediate our ability to distinguish textures. Furthermore, they contract or dilate systematically in time with increases or decreases in scanning speed. Thus, whereas coarse surface features are encoded in a spatial pattern of afferent activation, the fine spatial structure of a surface is converted into a temporal spiking pattern.

Spatial and temporal mechanisms are further elaborated at the cortical level. One population of neurons in somatosensory cortex is sensitive to the coarse spatial structure of skin deformations, on the order of millimeters. These neurons have receptive fields that contain excitatory and inhibitory subfields, analogous to those observed in primary visual cortex (DiCarlo et al., 1998). They are thought to perform spatial computations on the afferent

input, for example, explicitly encoding stimulus edges (Bensmaia et al., 2008), and likely contribute to the processing of coarse textural features. A separate population of somatosensory cortical neurons uses a time-varying firing rate to encode the envelope (time-varying amplitude) of skin vibrations; within this envelope, the fine structure of their firing patterns encodes the frequency composition of the vibrations (Harvey et al., 2013). The spectral composition of these spiking patterns shapes the way the vibrations (and by extension textures) are perceived. Judgments of roughness, the dominant dimension of texture, are well predicted based on a combination of spatial and temporal codes (Weber et al., 2013), suggesting that temporal representations of fine spatial features are combined with spatial representations of coarse features to culminate in a holistic percept of surface textures.

### High-dimensional computational architecture of odor representations

In olfaction, the dimensionality and geometry of similarity for representational processing are unusually configured, and this has important implications for the architecture of the olfactory system. In olfaction, dimensionality is unavoidably high and the geometry lacks a simple, externally defined basis analogous to wavelength or pitch. This substantially constrains effective neural algorithms and processing compared with early vision and audition. For example, the high dimensionality of the input space renders low-dimensional contrast enhancement algorithms, such as nearest-neighbor lateral inhibition, ineffective for olfaction, whereas the dependence of receptive field similarities on the statistical distribution of odorous inputs limits the efficacy of any scheme dependent on durable intercolumnar interactions. These problems have been resolved theoretically by a nontopographical algorithm based upon intracolumnar feedforward inhibition and global normalization within the glomerular layer (Cleland and Sethupathy, 2006; Cleland et al., 2007; Cleland, 2010; Li and Cleland, 2013) that is receiving increased experimental support for its predictions (Soucy et al., 2009; Gire and Schoppa, 2009; Marbach and Albeanu, 2011; D’Souza and Vijayaraghavan, 2012).

This, however, is only the first step. Sensory systems do not sample their input spaces uniformly; rather, they devote disproportionate resources to ecologically important regions of the perceptual space. For example, the cochlear nucleus in humans exhibits a substantial expansion in the regions devoted to speech frequencies. The Penfield somatosensory homunculus depicts the deeply nonuniform distribution of somatosensory receptors across the skin, heavily favoring the hands and face. Terrestrial visual systems appear to be optimized for the  $1/f$  scene statistics of spatial frequency exhibited by the natural world. The olfactory system is likely to exhibit such adaptive nonuniformities as well; however, the statistics of important environmental odors appear to be less predictable, compared with the size of the space, than the statistics of other modalities. Cleland and colleagues (Cleland, 2010; Cleland and Sethupathy, 2006; Cleland et al., 2007) have hypothesized that this function is achieved in olfaction by learning. Specifically, they propose that learning processes within olfactory bulb, probably in correspondence with piriform cortical feedback, serve to transform this vast high-dimensional space into probabilistic categorical representations based upon individual olfactory experience. These learned nonuniformities will profoundly bias the internal representational space in accordance with the evolving needs of the organism. The mechanisms proposed to underlie this adaptive transformation generate higher-order receptive fields in bulbar interneurons and cortical

principal neurons that enable near-arbitrary levels of separation among the representations of learned odors. Interestingly, these proposed mechanisms depend on the coordinated timing of action potentials rather than on differences in aggregate spike rates.

The geometry of odor similarity is quantifiable both in terms of perceptual similarity (behavior) and in terms of neuronal activation patterns: from primary olfactory receptors to higher-order principal neurons. Briefly, a given species expresses a species-specific complement of odorant receptors (ORs). ORs respond strongly to preferred odors, but also to “neighboring” odors, including several structural and side-chain variants (Araneda et al., 2000). These arbitrary but consistent receptive fields of ORs form the physical foundation of olfactory perceptual space. Unlike retinal cones or auditory hair cells, however, they lack any external basis for quantitative similarity other than the probability of coactivation. The most important consequence of this fact is that the similarity between the receptive fields of two ORs is not fixed but depends upon the statistical structure of the odor environment. (For example, two ORs may have essentially identical receptive fields, and be functionally 100% redundant, until the day comes when a new odor is encountered that activates one OR and not the other). This also is the reason why the receptive fields of different ORs must be considered independent in this context, such that the full dimensionality of odor space equals at least the number of different OR types. In any given, finite universe of odorants, the dimensionality of odor space can be reduced, often substantially; however, this lowered dimensionality reflects primarily the relative poverty of the input space. If the olfactory system is to be able to interpret any possible combination of receptor inputs received, then a high-dimensional input space is unavoidable, and the neural circuitry of the system must accommodate this fact.

## Discussion

This paper is not meant to be a comprehensive review of work on perceptual spaces, so it does not attempt to cover perceptual spaces that have been proposed for many other attributes, such as gloss (Ferwerda et al., 2001; Wills et al., 2009), timbre (Lakatos, 2000; Terasawa et al., 2005), vowels (Pols et al., 1969), gestures (Arfib et al., 2002), and biological motion (Giese and Lappe, 2002), mostly on the basis of multidimensional scaling analyses. Instead, it highlights six quite diverse approaches to some common central themes. One central theme concerns the perceived similarity among stimuli. At an abstract level, similarity can be used to generate models of generalization ranging from set-theoretic (Tversky, 1977) to continuous metric-space structures (Shepard, 1987), notably within a Bayesian formulation (Tenenbaum and Griffiths, 2001). Research on color mid-point settings and odor generalization provides examples of using relative similarity judgments to investigate the geometry of perceptual spaces. Such investigations are logical preliminaries to using multidimensional scaling to define a perceptual space, as they test the validity of the assumptions inherent in the statistical procedure. A second central theme is the efficient representation of environmentally important properties of natural stimuli. Two interesting aspects of such representations are exhibited across multiple sensory modalities. First, invariant representations may be contained within the statistics of distributions of stimuli. One approach to identify these invariances applies statistics of different orders to white noise to synthesize stimuli that may be metamers with real stimuli; another looks for the distribution of specific local features within stimuli. Second, across several modalities, audition, somatosensation, olfaction, and gustation (Di

Lorenzo et al., 2009), the representation of fine-grained properties exploits temporal spiking patterns and not just rates of firing. Indeed, correlated spiking patterns provide a robust substrate for invariant representations of complex stimuli. These analyses together help to demonstrate how the structure of model perceptual spaces represents a trade-off between the goals of providing an efficient representation of sensory stimuli and the constraints of neural hardware.

## References

- Araneda RC, Kini AD, Firestein S (2000) The molecular receptive range of an odorant receptor. *Nat Neurosci* 3:1248–1255. [CrossRef Medline](#)
- Arfib D, Couturier JM, Kessouf L, Verfaillie V (2002) Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Org Sound* 7: 2:127–144.
- Bensmaia SJ, Hollins M (2003) The vibrations of texture. *Somatosens Mot Res* 20:33–43. [CrossRef Medline](#)
- Bensmaia SJ, Denchev PV, Dammann JF 3rd, Craig JC, Hsiao SS (2008) The representation of stimulus orientation in the early stages of somatosensory processing. *J Neurosci* 28:776–786. [CrossRef Medline](#)
- Brannan DA, Esplen MF, Gray JJ (1999) *Geometry*. Cambridge, MA: Cambridge UP.
- Bregman AS (1990) *Auditory scene analysis: the perceptual organization of sound*. Cambridge, MA: MIT.
- Briguglio J, Hermundstad A, Conte MM, Victor JD, Tkačik G, Balasubramanian V (2013) Perceptual salience of fourth-order visual textures and natural scene statistics. *J Vis* 13:1234. [CrossRef](#)
- Carruthers IM, Natan RN, Geffen MN (2013) Encoding of ultra-sonic vocalizations in the rat auditory cortex. *J Neurophysiol*. 109:1912–1927. [CrossRef Medline](#)
- Cleland TA (2010) Early transformations in odor representation. *Trends Neurosci* 33:130–139. [CrossRef Medline](#)
- Cleland TA, Sethupathy P (2006) Non-topographical contrast enhancement in the olfactory bulb. *BMC Neurosci* 7:7. [CrossRef Medline](#)
- Cleland TA, Johnson BA, Leon M, Linster C (2007) Relational representation in the olfactory system. *Proc Natl Acad Sci U S A* 104:1953–1958. [CrossRef Medline](#)
- Connor CE, Johnson KO (1992) Neural coding of tactile texture: comparison of spatial and temporal mechanisms for roughness perception. *J Neurosci* 12:3414–3426. [Medline](#)
- Coxeter HSM, Greitzer SL (1967) Quadrangles: Varignon’s theorem. In: *Geometry revisited*, pp 51–56. Washington, DC: Mathematics Association of America.
- DiCarlo JJ, Johnson KO, Hsiao SS (1998) Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey. *J Neurosci* 18:2626–2645. [Medline](#)
- Di Lorenzo P, Chen JY, Victor JD (2009) Quality time: representation of a multidimensional sensory domain through temporal coding. *J Neurosci* 29:9227–9238. [CrossRef Medline](#)
- D’Souza RD, Vijayaraghavan S (2012) Nicotinic receptor-mediated filtering of mitral cell responses to olfactory nerve inputs involves the  $\alpha 3\beta 4$  subtype. *J Neurosci* 32:3261–3266. [CrossRef Medline](#)
- Ennis R, Zaidi Q (2013) Geometrical structure of perceptual color space is affine. *J Vis* 13:295. [CrossRef](#)
- Ferwerda JA, Pellacini F, Greenberg D (2001) A psychophysically-based model of surface gloss perception. *Proc SPIE* 4299:291–301. [CrossRef](#)
- Freeman J, Simoncelli EP (2011) Metamers of the ventral stream. *Nat Neurosci* 14:1195–1201. [CrossRef Medline](#)
- Gaver WW (1993) How do we hear in the world? Explorations in ecological acoustics. *Ecol Psychol* 5:285–313. [CrossRef](#)
- Giese MA, Lappe M (2002) Measurement of generalization fields for the recognition of biological motion. *Vision Res* 42:1847–1858. [CrossRef Medline](#)
- Gire DH, Schoppa NE (2009) Control of on/off glomerular signaling by a local GABAergic microcircuit in the olfactory bulb. *J Neurosci* 29:13454–13464. [CrossRef Medline](#)
- Harvey MA, Saal HP, Dammann JF 3rd, Bensmaia SJ (2013) Multiplexing stimulus information through rate and temporal codes in primate somatosensory cortex. *PLoS Biol* 11:e1001558. [CrossRef Medline](#)
- Hollins M, Bensmaia S, Karlof K, Young F (2000) Individual differences in

- perceptual space for tactile textures: evidence from multidimensional scaling. *Percept Psychophys* 62:1534–1544. [CrossRef Medline](#)
- Hollins M, Bensmaia SJ, Roy EA (2002) Vibrotactile and texture perception. *Behav Brain Res* 135:51–56. [CrossRef Medline](#)
- Isabelle SK, Colburn HS (1991) Detection of tones in reproducible narrow-band noise. *J Acoust Soc Am* 89:352–359. [CrossRef Medline](#)
- Klein F (1939) Elementary mathematics from an advanced standpoint. In: *Geometry, 3: Systematic discussion of geometry and its foundations*. Mineola, NY: Dover.
- Lakatos S (2000) A common perceptual space for harmonic and percussive timbres. *Percept Psychophys* 62:1426–1439. [CrossRef Medline](#)
- Li G, Cleland TA (2013) A two-layer biophysical model of cholinergic neuromodulation in olfactory bulb. *J Neurosci* 33:3037–3058. [CrossRef Medline](#)
- Mackevicius EL, Best MD, Saal HP, Bensmaia SJ (2012) Millisecond precision spike timing shapes tactile perception. *J Neurosci* 32:15309–15317. [CrossRef Medline](#)
- MacLeod DI, Boynton RM (1979) Chromaticity diagram showing cone excitation by stimuli of equal luminance. *J Opt Soc Am* 69:1183–1186. [CrossRef Medline](#)
- Marbach F, Albeanu D (2011) Photostimulation of short axon cells reveals widespread inhibition in the mouse olfactory bulb. *Soc Neurosci Abstr* 475:19.
- Maxwell JC (1860) On the theory of compound colours and the relations of the colours of the spectrum. *Phil Trans R Soc* 150.
- McDermott JH, Simoncelli EP (2011) Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. *Neuron* 71:926–940. [CrossRef Medline](#)
- McDermott JH, Oxenham AJ, Simoncelli EP (2009) Sound texture synthesis via filter statistics. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. New Paltz.
- McDermott JH, Schemitsch M, Simoncelli EP (2013) Summary statistics in auditory perception. *Nat Neurosci* 16:493–498. [CrossRef Medline](#)
- Nelken I, Bar-Yosef O (2008) Neurons and objects: the case of auditory cortex. *Front Neurosci* 2:107–113. [CrossRef Medline](#)
- Pfafflin SM, Mathews MV (1966) Detection of auditory signals in reproducible noise. *J Acoust Soc Am* 39:340–345. [CrossRef Medline](#)
- Pinto N, Doukhan D, DiCarlo JJ, Cox DD (2009) A high-throughput screening approach to discovering good forms of biologically inspired visual representation. *PLoS Comput Biol* 5:e1000579. [CrossRef Medline](#)
- Pols LC, van der Kamp LJ, Plomp R (1969) Perceptual and physical space of vowel sounds. *J Acoust Soc Am* 46:458–467. [CrossRef Medline](#)
- Portilla J, Simoncelli EP (2000) A parametric texture model based on joint statistics of complex wavelet coefficients. *Int J Comput Vis* 40:49–71. [CrossRef](#)
- Risset JC, Wessel DL (1999) Exploration of timbre by analysis and synthesis. In: *The Psychology of music* (Deutsch D, ed), pp 113–169. San Diego: Academic.
- Russ BE, Lee YS, Cohen YE (2007) Neural and behavioral correlates of auditory categorization. *Hear Res* 229:204–212. [CrossRef Medline](#)
- Schnapf JL, Kraft TW, Baylor DA (1987) Spectral sensitivity of human cone photoreceptors. *Nature* 325:439–441. [CrossRef Medline](#)
- Shepard RN (1962) The analysis of proximities: multidimensional scaling with an unknown distance function. Part I. *Psychometrika* 27:125–140. [CrossRef](#)
- Shepard RN (1987) Toward a universal law of generalization for psychological science. *Science* 237:1317–1323. [CrossRef Medline](#)
- Smith VC, Pokorny J (1975) Spectral sensitivity of the foveal cone photopigments between 400 nm and 500 nm. *Vision Res* 15:161–171. [CrossRef Medline](#)
- Soucy ER, Albeanu DF, Fantana AL, Murthy VN, Meister M (2009) Precision and diversity in an odor map on the olfactory bulb. *Nat Neurosci* 12:210–220. [CrossRef Medline](#)
- Tenenbaum JB, Griffiths TL (2001) Generalization, similarity, and Bayesian inference. *Behav Brain Sci* 24:629–640; discussion 652–791. [Medline](#)
- Terasawa H, Slaney M, Berger J (2005) The thirteen colors of timbre. *Proc IEEE WASPAA*, Mohonk, NY.
- Tversky A (1977) Features of similarity. *Psychol Rev* 84:327–352. [CrossRef](#)
- Victor JD, Conte MM (2012) Local image statistics: maximum-entropy constructions and perceptual salience. *J Opt Soc Am A* 29:1313–1345. [CrossRef Medline](#)
- Weber AI, Saal HP, Lieber JD, Cheng JW, Manfredi LR, Dammann JF III, Bensmaia SJ (2013) Spatial and temporal codes mediate the tactile perception of natural textures. *Proc Natl Acad Sci U S A* 110:17107–17112. [CrossRef Medline](#)
- Wertheimer M (1912) Experimentelle Studien über das Sehen von Bewegung [Experimental studies on the seeing of motion]. *Z Psychol* 61:161–265.
- Wills J, Agarwal S, Kriegman D, Belongie S (2009) Toward a perceptual space for gloss. *ACM Trans* 28:103.
- Wuerger SM, Maloney LT, Krauskopf J (1995) Proximity judgments in color space: tests of a Euclidean color geometry. *Vision Res* 35:827–835. [CrossRef Medline](#)
- Zaidi Q (1998) Identification of illuminant and object colors: heuristic-based algorithms. *J Opt Soc Am A* 15:1767–1776. [CrossRef Medline](#)
- Zaidi Q, Bostic M (2008) Color strategies for object identification. *Vision Res* 48:2673–2681. [CrossRef Medline](#)